PROBABILITY THEORY II

BY

TANUJIT CHAKRABORTY

Indian Statistical Institute

Mail : tanujitisi@gmail.com

PROBABILITY

" It is a measure of chance of occurrance of a Phenomenon."

NAME:- TANUJIT CHAKRABORTY.
TOPIC:- PROBABILITY THEORY 2.
HONS:- STATISTICS
CLG:- BIDHANNAGAR COLLEGE
YEAR:- B.SC 2ND YEAR

The problem in Probability is —
" Given a stochastic model, what we can say about the outcome".

" STATISTICS "

Spores of Information → $\boxed{\text{Statistical Method}}$ → A valid Decision.

PROBABILITY THEORY 2.

Generating Function :- the generating function of a random variable X, is a function of the form $E(\psi(t,X))$; where t is a non-random variable.

- **P.g.f. (Probability Generating Function):-** This is meant for a discrete random variable whose mass points are non-negative integers or some subsets of the whole set of non-negative integers. Here $\psi(t,x)$ is of the form t^X. Note that $E(t^X)$ necessarily exists for $|t|=1$. Hence, because of the comparison test we find that the series $p_0 + t.p_1 + t^2 p_2 + \ldots\ldots$ is also absolutely convergent for $|t|<1$. As such the p.g.f. of a non-negative integer valued random variable necessarily exists. It is denoted by $P_X(t)$.

Example:

1. Binomial Distribution :- (with parameter n, p)

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad ; \quad x \geq 0$$

$$P(t) = E(t^X)$$

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

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

$$= (q+pt)^n \text{, defined for all real } t.$$

2. Poisson Distribution :- (with parameter λ)

$$f(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} \quad , x \geq 0 \, ,$$

$$P(t) = E(t^X) = \sum_{x=0}^{\infty} t^x e^{-\lambda} \cdot \frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} e^{-\lambda} \frac{(\lambda t)^x}{x!}$$

$$= e^{-\lambda} \sum_{x} \frac{(\lambda t)^x}{x!}$$

$$= e^{\lambda t} \cdot e^{-\lambda}$$

$$= e^{-\lambda(1-t)} = e^{\lambda(t-1)}.$$

3. Negative Binomial Distribution :- (with parameter r, p)

$$f(x) = \binom{x+r-1}{x} p^r q^x \;;\; x = 0, 1, \ldots \text{ ad inf}$$

$$P(t) = E(t^X) = \sum_{x=0}^{\infty} t^x \binom{x+r-1}{x} p^r q^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} (qt)^x$$

$$= p^r (1-qt)^{-r}, \text{ defined for } |qt| < 1$$

(if the series is absolutely convergent) $|t| < \frac{1}{q}$.

▨ USES :- 1. As the name suggests a p.g.f. gives the probabilities corresponding to the mass points of the relevant discrete distribution, i.e. the p.m.f. of a discrete distn. The following theorem is useful in this respect:

● Theorem :-

If X, Y are non-negative integer valued r.v.s with p.g.f. $P_X(t)$ and $P_Y(t)$ then, provided X and Y are independent, the p.g.f. of $(X+Y)$ is $P_X(t) P_Y(t)$.

Proof :- The p.g.f of X, $P_X(t)$ be defined for $|t| < t_1$, say and the p.g.f of Y, $P_Y(t)$ be defined for $|t| < t_2$.

Let us take, $t_0 = \min\{t_1, t_2\}$

Then, we have for $|t| < t_0$,

$P_{X+Y}(t) = E(t^{X+Y})$ is defined

$$E(t^{X+Y}) = E(t^X . t^Y)$$

$$= E(t^X) . E(t^Y) \quad [\because X \text{ and } Y \text{ are independent, } t^X \,\&\, t^Y \text{ " " }]$$

$$= P_X(t) . P_Y(t) . \quad [\text{By product law of expectation}]$$

▨ Example :- Let ~~X has the Binom~~ $X \sim Bin(n_1, p)$ & $Y \sim Bin(n_2, p)$ & they are independent. Then find the probability distn. of $(X+Y)$.

ANS :

$$P_X(t) = (q + pt)^{n_1}$$

$$\& \; P_Y(t) = (q + pt)^{n_2}, \text{ defined for all } t.$$

Since X and Y are independent,

$$P_{X+Y}(t) = P_X(t) . P_Y(t)$$

$$= (q+pt)^{n_1+n_2}, \text{ defined for all } t.$$

which is itself the p.g.f. of another Binomial distrn with parameter (n_1+n_2, p). $\therefore X+Y \sim Bin(n_1+n_2, p)$

2. Suppose the moments of all order upto r of a non-negative integer valued random variable X exists. Then the rth factorial moment of X can be obtained from $P_X(t)$ by differentiation.

$$P_X(t) = E(t^X) = \sum_{x=0}^{\infty} t^x p_x$$

$$\frac{d^r}{dt^r} P_X(t) = \sum_{x=0}^{\infty} x(x-1)(x-2)\cdots(x-r+1) \cdot t^{x-r} \cdot p_x$$

$$\left.\frac{d^r}{dt^r} P_X(t)\right|_{t=1} = \sum_{x=0}^{\infty} x_{[r]}\, p_x = E[X_{[r]}]$$

Particular case:-

$$P'_X(1) = E(X) \quad \forall t=1$$

$$P''_X(1) = E[X(X-1)] = E(X^2) - E(X)$$

$$\therefore P''_X(1) + P'_X(1) = E(X^2)$$

$$\therefore P''_X(1) + P'_X(1) - \{P'_X(1)\}^2 = Var(X)$$

☆ 8.1. Find the p.g.f. of the no. of points to be obtained in throwing a fair die once. Also find the p.g.f. of the no. of points to be obtained in throwing a fair die n times. Hence s.t. the Prob. of obtaining s points in n times is the prob. of obtaining $(7n-s)$ points.

ANS:- i)

$X : 1, 2, 3, 4, 5, 6$.

$X =$ Point obtained in one throw.

$$P_X(t) = E(t^X) = \sum_{x=1}^{6} t^x . P(X=x)$$

$$= \frac{1}{6}(t + t^2 + \cdots + t^6)$$

ii) $S =$ Total points obtained in n throws

$S = X_1 + X_2 + \cdots + X_n$; Note that X_i's are indep. r.v.'s, with the same distribution

$\therefore$ P.g.f. of S is

$P(X_i = x) = \frac{1}{6} \; \forall x = 1(1)6$, $i = 1(1)n$.

Hence, each of them has the same p.g.f.

$$P_S(t) = E(t^S) = E(t^{X_1 + X_2 + \cdots + X_n})$$

[$\because X_i$'s are indep.] $= E(t^{X_1}) E(t^{X_2}) \cdots E(X^n) = \{P_X(t)\}^n = \frac{1}{6^n}(t+t^2+\cdots+t^6)^n$

iii) $P(7n-s \text{ points in } n \text{ throws})$

$= \text{coefficient of } t^{7n-s} \text{ in } \frac{1}{6^n}(t+t^2+\cdots+t^6)^n$.

$= \text{coefficient of } t^{-s} \text{ in } \frac{1}{6^n}(t^{-1}+t^{-2}+\cdots+t^{-6})^n$

$= \text{coefficient of } t^{s} \text{ in } \frac{1}{6^n}(t+t^2+\cdots+t^6)^n$

$= P(s \text{ points in } n \text{ throws})$.

☆ 8.4. If X is a r.v. assuming integral non-negative values and with p.g.f $P_X(t)$, then s.t. for $0<a<b$

i) $E\left(\frac{1}{X+a}\right) = \int_0^1 t^{a-1} P_X(t)\,dt$

ii) $E\left(\frac{1}{(X+a)(X+b)}\right) = \int_0^1 t^{b-a-1} \int_0^t u^{a-1} P_X(u)\,du\,dt$

ANS:-

i) $\int_0^1 t^{a-1} P_X(t)\,dt$

$= \int_0^1 t^{a-1} \left(\sum_x t^x p_x\right) dt$

$= \sum_x \left[\int_0^1 t^{x+a-1}\,dt\right] p_x$ [the interchange of summation and integration is possible as sum is finite]

$= \sum_x \left[\frac{t^{x+a}}{x+a}\right]_0^1 p_x$

$= \sum_x \frac{1}{x+a} p_x$

$= E\left(\frac{1}{X+a}\right)$

ii) $\int_0^1 t^{b-a-1} \int_0^t u^{a-1} P_X(u)\,du\,dt$

$= \int_0^1 t^{b-a-1} \int_0^t \left[u^{a-1}\left(\sum_x u^x p_x\right)\right] du\,dt$

$= \sum_x \left[\int_0^1 t^{b-a-1}\left(\int_0^t u^{a-1} u^x\,du\right) dt\right] p_x$

$= \sum_x \left[\int_0^1 t^{b-a-1}\left(\int_0^t u^{a+x-1}\,du\right) dt\right] p_x$

$= \sum_x \left[\int_0^1 t^{b-a-1}\left(\frac{u^{a+x}}{a+x}\right)_0^t dt\right] p_x$

$= \sum_x \left[\int_0^1 t^{b-a-1}\,\frac{t^{x+a}}{x+a}\,dt\right] p_x$

$$= \sum_x \left[\int_0^1 \frac{t^{x+b-1}}{x+a} dt \right] p_x$$

$$= \sum_x \left[\frac{t^{x+b}}{(x+a)(x+b)} \right]_0^1 p_x$$

$$= \sum_x \frac{1}{(x+a)(x+b)} p_x$$

$$= E\left(\frac{1}{(X+a)(X+b)} \right)$$

☆ 8.5. Let X be a geometrically distributed R.V. with p.m.f.

$$f(x) = pq^x \; ; \; x = 0, 1, 2, \ldots$$
$$= 0 \quad , \text{o/w}$$

where $0 < p < 1$; $q = 1 - p$.

Find the p.g.f. of X and those of $X_+ = \text{Max}[N, X]$

$X_- = \text{Min}[N, X]$

for a given positive integer N.

ANS: –

The p.g.f. is $P_X(t) = \sum_{x=0}^{\infty} t^x P(X = x)$

$$= \sum_{x=0}^{\infty} t^x . pq^x$$

$$= \sum_{x=0}^{\infty} p(qt)^x$$

$$= \frac{p}{1-qt}$$ [Expansion valid for $|qt| < 1$ or, $|t| < \frac{1}{q}$]

$\therefore$

To find p.g.f.s of X^+ & X^-, need to find $P(X^+ = x)$ & $P[X^- = x]$

Now, $P(X^+ = x) = P(X \leq N)$ $\quad$ $X > N$, $X^+ = X$

$$= \sum_{x=0}^{N} pq^x$$ $\quad$ $X < N$, $X^+ = N$

$$= \frac{p(1 - q^{N+1})}{1 - q}.$$

$$= 1 - q^{N+1}$$

and $P(X_+ = x) = P(X = x)$ for $x > N$

$= pq^x$ for $x = N+1, N+2, \ldots\ldots$ inf

$\therefore P(X_+ = x) = 1 - q^{N+1}$ if $x = N$

$= pq^x$ if $x > N$

p.g.f. of $X+$

$$P_{X+}(t) = \sum_{x=N}^{\infty} t^x P(X_+ = x)$$

$$= t^N(1-q^{N+1}) + \sum_{x=N+1}^{\infty} t^x pq^x$$

$$= t^N(1-q^{N+1}) + p(qt)^{N+1} \cdot \frac{1}{(1-qt)}$$ defined for $|qt|<1$, $|t|<\frac{1}{q}$

(Ans)

Now, $P[X^- = x] = P[X \geq N]$ $X \geq N$, $X^- = N$; $X<N$, $X^- = X$

$$= 1 - P[X \leq N]$$

$$= 1 - (1-q^{N+1})$$

$$= q^{N+1}$$

and

$$P(X_- = x) = P(X = x) \text{ for } x < N$$

$$= pq^x \text{ for } x = 0, 1, \ldots, N-1$$

$$\therefore P(X_- = x) = q^{N+1} \text{ if } X = N$$

$$= pq^x \text{ if } X < N$$

Now, p.g.f of $X-$

$$P_{X-}(t) = \sum_{x=0}^{N} t^x P(X_- = x)$$

$$= t^N q^{N+1} + \sum_{x=0}^{N-1} t^x pq^x$$

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

$$= t^N q^{N+1} + \frac{p(1-(qt)^N)}{(1-qt)}$$ defined for $|qt|<1$

$\Rightarrow |t| < \frac{1}{q}$.

i.e. $t \neq \frac{1}{q}$.

☆ **8.6.** Find the p.g.f.s of the following distributions:

(a) $P[X=x] = \dfrac{e^{-\lambda}\lambda^{x}}{(1-e^{-\lambda})x!}$ for $x = 1, 2, \ldots\ldots$, where $\lambda > 0$;

(b) $P[X=x] = pq^{x}(1-q^{N+1})^{-1}$ for $x = 0, 1, \ldots, N$, where $0<p<1$, $q = 1-p$.

ANS: —

(a)
$$P_X(t) = E(t^X) = \sum_{x=1}^{\infty} t^x P[X=x]$$
$$= \sum_{x=1}^{\infty} t^x \cdot \frac{e^{-\lambda}\cdot\lambda^x}{(1-e^{-\lambda})x!}$$
$$= \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{x=1}^{\infty} \frac{(\lambda t)^x}{x!}$$
$$= \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(\sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} - 1 \right)$$
$$= \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(e^{\lambda t} - 1\right) \quad \text{[ANS]}$$

(b)
$$P_X(t) = E(t^X) = \sum_{x=0}^{N} t^x P[X=x]$$
$$= \sum_{x=0}^{N} t^x \cdot \frac{pq^x}{(1-q^{N+1})}$$
$$= \frac{p}{(1-q^{N+1})} \sum_{x=0}^{N} t^x q^x$$
$$= \frac{p}{(1-q^{N+1})} \sum_{x=0}^{N} (qt)^x \quad \text{[finite g.p. series]}$$
$$= \frac{p}{(1-q^{N+1})} \cdot \frac{1\left(1-(qt)^{N+1}\right)}{(1-qt)}$$
$$= \frac{p\left(1-(qt)^{N+1}\right)}{(1-q^{N+1})(1-qt)} \quad \text{[ANS]}$$

▨ **NOTE:→** Every pgf determines a unique probability i.e. a unique set of probabilities $\{p_x\}$. Coefficient of t^x in the expansion of $P_X(t)$ gives $P(X=x)$. The pgf is used for discrete variables only. Probability for a particular point can only be determined for discrete random variables. So, pgf can't be determined for continuous variables.

☆ **8.7.** Let $P_1(t)$ and $P_2(t)$ be the p.g.f.s of the two independently distributed r.v.s X_1 and X_2 assuming non-negative integral values. Show that $P[X_1 - X_2 = j]$ is the coefficient of t^j in $P_1(t) \cdot P_2\left(\frac{1}{t}\right)$, where $j = 0, \mp 1, \mp 2, \dots\dots$.

ANS:- Need to show, $P_{X_1-X_2}(t) = P_1(t) \cdot P_2\left(\frac{1}{t}\right)$

$$P_{X_1-X_2}(t) = E\left(t^{X_1-X_2}\right) = E\left(t^{x_1} \cdot t^{-x_2}\right)$$

$$= E\left(t^{X_1}\right) \cdot E\left(t^{-x_2}\right) \quad [\because X_1, X_2 \text{ ind.}]$$

$$= E\left(t^{X_1}\right) \cdot E\left(\left(\frac{1}{t}\right)^{X_2}\right)$$

$$= P_1(t) \cdot P_2\left(\frac{1}{t}\right)$$

☆ **8.19.** Let $\{X_n\}$ be a sequence of i.i.d. random variables with the common distribution given by $P[X_n = x] = f_x$ and p.g.f. $f(t) = \sum f_x t^x$. If $S_N = X_1 + X_2 + \dots + X_N$ and N is itself a random variable with $P[N=n] = g_n$ and has the p.g.f. $g(t) = \sum g_n t^n$, obtain the p.g.f. of S_N.

ANS:- P.g.f. of S_N

$$P_{S_N}(t) = E\left(t^{S_N}\right)$$

$$= E_N\left[E\left(t^{S_N} \mid N=n\right)\right]$$

$$= E_N\left[E\left(t^{S_N}\right)\right]$$

$$= E_N\left[E\left(t^{X_1+\dots+X_n}\right)\right]$$

$$= E_N\left[\prod_{i=1}^{n} E\left(t^{X_i}\right)\right] \quad [\because X_i\text{'s are i.i.d. with p.g.f } f(t)]$$

So, $P_{S_N}(t) = E_N\left[\prod_{i=1}^{n} f(t)\right]$

$= E_N\left[(f(t))^N\right]$

$= \sum_n (f(t))^n P(N=n)$

$= \sum_n (f(t))^n g_n$

$= g(f(t))$

☆ 8.2.D. (Continuation) If X_n $(n=1,2,\ldots)$ be independent Bernoulli random variables with parameter p and N is a Poisson r.v. with parameter λ, obtain the p.g.f. of S_N. Show that S_N is itself a Poisson random variable having parameter λp.

ANS:- $f_x = P(X_n = x) = p^x q^{1-x}$; $x = 0, 1$.

$g_n = P(N=n) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$

$\therefore g(t) = e^{-\lambda(1-t)}$

$f(t) = (q+tp)$.

$\therefore P_{S_N}(t) = g(f(t))$

$= g(q+pt)$

$= e^{-\lambda(1-q-pt)}$

$= e^{-\lambda(p-pt)}$

$= e^{-\lambda p(1-t)}$ $\rightarrow$ It is the p.g.f. of a Poisson r.v. with parameter λp.

☆ 8.21. (Continuation) If X_n ($n = 1, 2, \ldots$) are independently distributed each having the negative binomial distribution with parameters r and p, and N has the Poisson distribution with parameter λ, what is the p.g.f. of S_N?

ANS:-

$$f_x = P(X_n = x) = \binom{x+r-1}{r-1} p^r q^x$$

$$g_n = P(N = n) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

$$\therefore f(t) = p^r (1-qt)^{-r} \quad \text{for } |t| < \frac{1}{q}$$

$$\therefore g(t) = e^{-\lambda(1-t)}$$

$$\therefore P_{S_N}(t) = g(f(t)) = g\left(p^r(1-qt)^{-r}\right)$$

$$= e^{-\lambda(1 - p^r(1-qt)^{-r})}$$

$$= e^{[\lambda p^r(1-qt)^{-r} - \lambda]} \quad \text{for } |t| < \frac{1}{q}$$

[Ans]

Check:

$$P_{S_N}(t) = e^{\lambda(t-1)} \stackrel{?}{=} e^{\lambda(p^r(1-qt)^{-r} - 1)}$$

☆ 8.22. (Compound Poisson distribution) If N has the Poisson distribution with parameter λ while X_n ($n = 1, 2, \ldots$) are independently distributed each having the p.g.f. $f(t)$, obtain the p.g.f. of $S_N = X_1 + X_2 + \cdots + X_N$.

ANS:-

$$P_{S_N}(t) = g(f(t)), \text{ where } g(t) = e^{\lambda(t-1)}.$$

$$= e^{\lambda\{f(t) - 1\}} \quad \text{(ANS)}$$

☆ 8.10. If X be a random variable taking non-negative integral values such that

$$P[X=j] = p_j, \quad \sum_{j=0}^{\infty} t^j p_j = P(t) \ \& \ q_j = P[X>j]$$

for $j = 0, 1, 2, \ldots\ldots$ write

$$Q(t) = \sum_{j=0}^{\infty} q_j t^j .$$

Show that $Q(t)$ converges for $|t|<1$. (It is called the generating function of $P[X>j]$. Show also that

$$Q(t) = \frac{1-P(t)}{1-t}, \quad |t|<1.$$

ANS:- Convergence :—

$$q_j < 1 \quad \forall j$$

$$\sum q_j t^j < \sum t^j, \quad \sum t^j \rightarrow \text{ converges absolutely for } |t|<1$$

By comparison test, the series $\sum_{j=0}^{\infty} q_j t^j$ also converges for $|t|<1$.

To prove,

$$Q(t) = \frac{1-P(t)}{1-t}, \quad |t|<1$$

$$Q(t) - tQ(t) = 1 - P(t) \quad \text{[To be proved]}$$

$$\text{L.H.S} = \sum_{j=0}^{\infty} q_j t^j - \sum_{j=0}^{\infty} q_j t^{j+1}$$

$q_0 = P(X>0)$
$= 1 - P(X=0)$
$= (1-p_0)$

$$= q_0 + t(q_1 - q_0) + t^2(q_2 - q_1) + \ldots\ldots$$

$(q_1 - q_0) = P(X>1) - P(X>0)$
$= -p_1$

$$= (1-p_0) + t(-p_1) + t^2(-p_2) + t^3(-p_3) + \ldots\ldots$$

$$= 1 - \sum_{j=0}^{\infty} t^j p_j$$

$$= 1 - P(t) = \text{R.H.S} \quad \text{(Proved)}$$

☆ 8.11. Find the generating functions of $P[X \le j]$, $P[X < j]$, $P[X \ge j]$, $P[X > j+1]$ and $P[X = 2j]$ in terms of $P(t)$.

ANS:— (i) $P(X \le j)$

Generating function of $P(X \le j)$ is for $|t| < 1$

$$Q(t) = \sum_{j=0}^{\infty} P(X \le j) t^j$$

$$= \sum_{j=0}^{\infty} \{ P(X = j) + P(X \le j-1) \} t^j$$

$$= \sum_{j=0}^{\infty} p_j t^j + \sum_{j=0}^{\infty} P(X \le j-1) t^j$$

$$= P(t) + t \sum_{j=0}^{\infty} P(X \le j-1) t^{j-1}$$

$$= P(t) + t\, Q(t)$$

$$Q(t) = \frac{P(t)}{1-t}$$

(ii) $P(X < j)$

Generating function of $P(X < j)$ is

$$Q_1(t) = \sum_{j=0}^{\infty} P(X < j) t^j$$

$$= \sum_{j=0}^{\infty} \{ P(X \le j) - P(X = j) \} t^j$$

$$= \sum_{j=0}^{\infty} P(X \le j) t^j - \sum_{j=0}^{\infty} P(X = j) t^j$$

$$= Q(t) - P(t)$$

$$= \frac{P(t)}{1-t} - P(t)$$

$$= \frac{t P(t)}{1-t}$$

(iii) $P(X \geq j)$

$$Q_2(t) = \sum_{j=0}^{\infty} P(X \geq j)\, t^j$$

$$= \sum_{j=0}^{\infty} \{1 - P(X < j)\}\, t^j$$

$$= \sum_{j=0}^{\infty} t^j - \sum_{j=0}^{\infty} P(X < j)\, t^j$$

$$= \frac{1}{1-t} - Q_1(t)$$

$$= \frac{1}{1-t} - \frac{tP(t)}{1-t}$$

$$= \frac{1 - tP(t)}{1-t}.$$

(iv) $P[X > j+1]$

$$Q_3(t) = \sum_{j=0}^{\infty} P[X > j+1]\, t^j$$

$$= \sum_{j=0}^{\infty} \{1 - P(X \leq j+1)\}\, t^j$$

$$= \sum_{j=0}^{\infty} t^j - \frac{1}{t} \sum_{j=0}^{\infty} P[X \leq j+1]\, t^{j+1}$$

$$= \frac{1}{1+t} - \frac{1}{t}\left(Q(t)\right) = \frac{1}{1+t} - \frac{P(t)}{t(1-t)}$$

(v) $P[X = 2j]$

$$Q_4(t) = \sum_{j=0}^{\infty} t^j P[X = 2j]$$

$$= \sum_{j=0}^{\infty} P\left[\frac{X}{2} = j\right] t^j$$

$$= E\left(t^{X/2}\right)$$

$$= P\left(t^{1/2}\right), \text{ for } t \geq 0.$$

☆ **8.13.** X & Y are two independent zero-truncated Poisson r.v.s. Find the p.g.f. of X+Y when
(a) X & Y are identically distributed with mean λ.
(b) X & Y have means λ and λ_2, respectively.

ANS:-
(a) $$P_{X+Y}(t) = \{P_X(t)\}^2 = [E(t^X)]^2 = \left[\sum_{x=1}^{\infty} t^x . P(X=x)\right]^2$$
$$= \left[\sum_{x=1}^{\infty} t^x . e^{-\lambda} . \frac{\lambda^x}{x!} (1-e^{-\lambda})^{-1}\right]^2 = \left[\frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{x=1}^{\infty} \frac{(\lambda t)^x}{x!}\right]^2$$
$$= \left\{\frac{e^{-\lambda}}{1-e^{-\lambda}} (e^{\lambda t}-1)\right\}^2$$
(b) $$P_{X+Y}(t) = P_X(t) P_Y(t) = \frac{e^{\lambda_1(t-1)} - e^{-\lambda_1}}{1-e^{-\lambda_1}} \times \frac{e^{\lambda_2(t-1)} - e^{-\lambda_2}}{1-e^{-\lambda_2}}$$

(Ans)

☆ **8.8.** Let X denote the number of trials required to get two consecutive heads when a fair coin is tossed. Show that the p.g.f. of X is

$$P(t) = \frac{t^2}{4}\left[1-\frac{t}{2} - \left(\frac{t}{2}\right)^2\right]^{-1}.$$

ANS:- Let p_n = P(In first n tosses, the sequence HH never occurs)
= P(Tail in 1st and then no HH in the next n−1)
+ P(Head in 1st and tell in 2nd and no HH in the next n−2)
$= \frac{1}{2} p_{n-1} + \frac{1}{4} p_{n-2}$ —— ① $\forall\ n \geq 2$

Let $\Phi(t)$ denote the p.g.f. of p_n; multipling ① by t^j and summing over $j=2,3,\ldots$

$$\sum_{j=0}^{\infty} t^j p_j = \sum_{j=2}^{\infty} t^j \left[\frac{1}{2} p_{j-1} + \frac{1}{4} p_{j-2}\right]$$
$$= \frac{1}{2} \sum_{j=2}^{\infty} t^j p_{j-1} + \frac{1}{4} \sum_{j=2}^{\infty} t^j p_{j-2}$$

$$\cancel{\frac{t}{2}\sum_{j=\bullet}^{\infty} t^{j-1} p_{j-1} + \frac{t^2}{4}\sum_{j=0}^{\infty} t^{j-2} p_{j-2}}$$

Now, $Q(t) = \sum_{j=0}^{\infty} t^j p_j$

$$Q(t) - tp_1 - p_0 = \frac{1}{2}\sum_{j=2}^{\infty} t^j p_{j-1} + \frac{1}{4}\sum_{j=2}^{\infty} t^j p_{j-2}$$

$$= \frac{1}{2}t\sum_{j=2}^{\infty} t^{j-1} p_{j-1} + \frac{t^2}{4}\sum_{j=2}^{\infty} t^{j-2} p_{j-2}$$

$$= \frac{1}{2}t\left(P(t) - p_0\right) + \frac{1}{4}(t^2)P(t)$$

Now, $p_0 = 1$, $p_1 = 1$. (obvious)

$$\Rightarrow Q(t) - t - 1 = \frac{1}{2}t\left(P(t) - 1\right) + \frac{1}{4}t^2 P(t)$$

$$\Rightarrow Q(t) - \frac{1}{2}tQ(t) - \frac{1}{4}t^2 Q(t) = t + 1 - \frac{1}{2}t$$

$$\Rightarrow Q(t)\left[1 - \frac{1}{2}t - \frac{1}{4}t^2\right] = \frac{1}{2}t + 1$$

$$\Rightarrow Q(t) = \frac{\frac{1}{2}t + 1}{\left(1 - \frac{1}{2}t - \frac{1}{4}t^2\right)}$$

X: No. of trials to get first HH.

$P(X=n) = \pi_n = P($In the nth toss first sequence HH occurs$)$

$$= p_{n-1} - p_n$$

Here, $\pi_0 = \pi_1 = 0$.

$\therefore$ p.g.f. of X is, $P(t) = \sum_{n=0}^{\infty} t^n \pi_n$, $|t| < 1$

$$= \sum_{n=2}^{\infty} t^n \pi_n$$

$$= \sum_{n=2}^{\infty} t^n (p_{n-1} - p_n)$$

$$= t\sum_{n=2}^{\infty} t^{n-1} p_{n-1} - \sum_{n=2}^{\infty} t^n p_n$$

$$= t\left[Q(t) - p_0\right] - \left[Q(t) - 1 - t\right]$$

$[\because p_0 = p_1 = 1]$

$$= t\,Q(t) - Q(t) + 1$$

$$= (t-1)Q(t) + 1$$

$$= \frac{t^2}{4}\left[1 - \frac{t}{2} - \left(\frac{t}{2}\right)^2\right]^{-1} = \frac{\frac{1}{2}(t+2)(t-1)}{\left(1 - \frac{1}{2}t - \frac{1}{4}t^2\right)} + 1$$

Probability Generating function in Bivariate Case :—

$\underset{\sim}{X} = (x_1, x_2)'$ be a bivariate random vector.
Then the p.g.f. of $\underset{\sim}{x}$ is defined as

$$P_{X_1,X_2}(t_1,t_2) = E\left(t_1^{X_1} \cdot t_2^{X_2}\right)$$

Let $f(x,y)$ denote the joint p.m.f. of (X, Y)

$$P_{X,Y}(t_1,t_2) = \sum_x \sum_y t_1^x t_2^y f(x,y)$$, provided the sum is absolutely convergent.

The p.g.f. of x_1 is obtained by putting $t_2 = 1$,

$$P_{X,Y}(t_1,1) = \sum_x \sum_y t_1^x f(x,y)$$
$$= \sum_x t_1^x \left[\sum_y f(x,y)\right]$$
$$= \sum_x t_1^x g(x)$$
$$= P_X(t_1).$$

☆ 8.23. The p.g.f. of the joint distribution of X_1 and X_2 is
$$P(t_1,t_2) = \exp\left[\lambda(p_1t_1 + p_2t_2 + p_3t_1t_2 - 1)\right],$$
where $p_i \geq 0 \ \forall i = 1,2,3$ and $p_1+p_2+p_3 = 1$.
(a) Find the marginal distribution of X_1 and X_2.
(b) Find the p.g.f. of the distribution of X_1+X_2.

ANS:-
(a) $P(t_1,t_2) = \exp\left[\lambda(p_1t_1 + p_2t_2 + p_3t_1t_2 - 1)\right]$; $p_i \geq 0$, $i = 2,3$. $\sum_1^3 p_i = 1$.

Marginal p.g.f. of x_1 is
$$P(t_1,1) = \exp\left[\lambda(p_1t_1 + p_2 + p_3t_1 - 1)\right]$$
$$= \exp\left[\lambda(t_1(p_1+p_3)) - (1-p_2)\right]$$
$$= \exp\left[\lambda(t_1(1-p_2)) - (1-p_2)\right]$$
$$= \exp\left[\lambda(1-p_2)(t_1-1)\right], \because p_1+p_2+p_3 = 1$$
which is p.g.f. of the Poisson Random variable with parameter $\lambda(1-p_2)$.
$\therefore X \sim \text{Poisson}(\lambda(1-p_2))$.

(b) P.g.f. of X_1+X_2 is $E\left(t^{X_1+X_2}\right)$

$$= E\left(t^{X_1}.t^{X_2}\right)$$

$$= P_{X_1,X_2}(t,t)$$

$$= \exp\left[\lambda\left(p_1 t + p_2 t + p_3 t^2 - 1\right)\right]$$

⇨ **Moment Generating function :—**

Here the form of $\psi(t,x) = e^{tx}$.

By the m.g.f. of a r.v. X, we mean $E(e^{tX})$, provided this expectation exists for all t satisfying $|t|<h$, i.e., $-h<t<h$, $h>0$

It is denoted by $M(t)$ or $M_X(t)$.

- **Example :—**

(i) Let $X \sim Bin(n,p)$, then $E(e^{tX})$ exists for all t,

Hence, the m.g.f. exists and is given by

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

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

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

$$= \left(pe^t+q\right)^n = \left(q+pe^t\right)^n$$

(ii) Let $X \sim Poiss(\lambda)$, the $E(e^{tX})$ exists for all t,

hence the m.g.f is defined and given by

$$M_X(t) = E\left(e^{tX}\right)$$

$$= \sum_{x=0}^{\infty} e^{tx}.\frac{e^{-\lambda}}{x!}.\lambda^x$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!}$$

$$= e^{-\lambda}.e^{\lambda e^t}$$

$$= e^{-\lambda(1-e^t)} = e^{\lambda(e^t-1)}$$

USES :-

1. Suppose all moments of X exists, then expanding $M_X(t)$ in a power series of X, we get, assume for all discrete X

$$M_X(t) = E(e^{tX})$$
$$= \sum_{x=0}^{\infty} e^{tx} \cdot P(X=x)$$
$$= \sum_{x=0}^{\infty} \left(\sum_{k=0}^{\infty} \frac{(tx)^k}{k!} \right) P(X=x)$$
$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{x=0}^{\infty} x^k P(X=x) \quad \text{, since the inner sum is finite}$$
$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot \mu_k'$$

Then the ~~rth~~ raw moment of X is obtainable as the co-efficient of $\frac{t^r}{r!}$ in the power series expansion of $M_X(t)$.

Hence, the name 'm.g.f.'.

The same thing can be shown for continuous case also.

Central moment generating function :-

Now, $M_{X-\mu}(t) = E\left(e^{t(X-\mu)}\right) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot \mu_k$. —— (1.1)

▨ The m.g.f. of a linear $Y = a+bX$ is

$$M_{a+bX}(t) = \int_{-\infty}^{\infty} e^{t(a+bx)}\, dF(x)$$
$$= e^{at} M_X(bt),$$

which is obtained by replacing t by bt in $M_X(t)$ and multiplying the result by e^{at}.

2. Let X and Y are independent random variables, with m.g.f. $M_X(t)$ and $M_Y(t)$ for t satisfying $-h_0 < t < h_0$ where $h_0 = \min\{h_1, h_2\}$ and $M_X(t)$ is defined for $-h_1 < t < h_1$. $M_Y(t)$ is defined for $-h_2 < t < h_2$.

Then the m.g.f. of $(X+Y)$ is given by

$$M_{X+Y}(t) = E\left(e^{t(X+Y)}\right), \quad |t| < h_0$$

$$= E\left(e^{tX} \cdot e^{tY}\right)$$

$$= E(e^{tX}) \cdot E(e^{tY}) \quad \left[\because X \,\&\, Y \text{ are indep. so are } e^{tX} \,\&\, e^{tY}\right]$$

$$= M_X(t)\, M_Y(t).$$

Defined for $-h_0 < t < h_0$.

Hence knowing the m.g.f. of X and Y, one can obtain the m.g.f. of $(X+Y)$, provided X and Y are independent.

Since the m.g.f. is unique, one can then obtain the distribution of $(X+Y)$.

- Example: Let $X \sim Bin(n_1, p)$, $Y \sim Bin(n_2, p)$ } independent.

$$\therefore \quad M_X(t) = (q + pe^t)^{n_1}$$

$$M_Y(t) = (q + pe^t)^{n_2}$$

$$\therefore M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

$$= (q + pe^t)^{n_1+n_2}$$

which is the m.g.f. of a $Bin(n_1+n_2, p)$ variable.

Hence, $X+Y \sim Bin(n_1+n_2, p)$

Find m.g.f.

(i) $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$

(ii) $f(x) = \theta e^{-\theta x}, \quad 0 \le x < \infty$

(iii) $f(x) = \dfrac{1}{B(m,n)} x^{m-1}(1-x)^{n-1}, \quad 0 < x < 1.$

3. The difficulty with m.g.f. is that it may not always exists. Suppose, x is a non-negative r.v. for which for some positive integer r $E(x^r)$ does not exist. Show that in this case, mgf does not exist.

$$e^{tx} = \sum_{r=0}^{\infty} \frac{(tx)^r}{r!}$$

for $t>0$, $e^{tx} > t^r . \frac{x^r}{r!}$

$$\Rightarrow \int_{-\infty}^{\infty} e^{tx} f(x)dx > \int_{-\infty}^{\infty} t^r . \frac{x^r}{r!} f(x)dx$$

$$= \frac{t^r}{r!} \int_{-\infty}^{\infty} x^r f(x)dx$$

since, $E(x^r)$ does not exist, $\int_{-\infty}^{\infty} x^r f(x)dx = \infty$

$\therefore \int_{-\infty}^{\infty} e^{tx} f(x)dx = \infty$. $\Rightarrow E(e^{tx})$ doesn't exist for $t>0$.

Hence the m.g.f. does not exist.

- Ex.:→ i) Cauchy distribution, $f(x) = \frac{1}{\pi} \frac{1}{(1+x^2)}$, $-\infty < x < \infty$,
  $\Rightarrow E(x)$ doesn't exist. $\Rightarrow$ MGF does not exist.

  ii) Lognormal distribution,
  MGF does not exist.

MGF in bivariate case :—

Let $\underset{\sim}{X} = (X_1, X_2)'$ is a bivariate random variable.
Then the MGF of $\underset{\sim}{X}$ is defined as

$$M_{X_1,X_2}(t_1, t_2) = E\left(e^{\underset{\sim}{t}'\underset{\sim}{X}}\right), \text{where } \underset{\sim}{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$$

Now, the m.g.f. of X_1 is obtained by putting $t_2 = 0$

$$M_{X_1,X_2}(t_1, 0) = E\left(e^{t_1 X_1}\right)$$

$$= M_{X_1}(t_1).$$

Note that, here, $M(t_1, t_2) = E\left(e^{t_1 X_i + t_2 X_j}\right)$ for all values of t_1 & t_2, is the moment generating function of the joint distribution of x_i & x_j.

$$E(X_i X_j) = \frac{\partial^2 M(0,0)}{\partial t_1 . \partial t_2}.$$

☆ Ques: - Find the moment generating function of the following distributions;

1. Exponential Distribution with parameter θ

The probability density function of X is given by

$$f(x) = \theta e^{-\theta x} \; ; \; x \geq 0$$

ANS:- $$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx}\, \theta e^{-\theta x}\, dx$$

$$= \theta \int_0^\infty e^{x(t-\theta)} dx \quad \text{, defined for } t < \theta$$

$$= \frac{\theta}{t-\theta} \int_0^\infty e^{z} dz$$

$$= \frac{\theta}{t-\theta} \left[e^{z} \right]_0^\infty$$

$$= \frac{\theta}{t-\theta}(-1) = \frac{\theta}{\theta - t}.$$

$x(t-\theta) = z$
$(t-\theta)dx = dz$
$\Rightarrow dx = \frac{dz}{(t-\theta)}$.

Using $M_X(t)$, finding mean & variance:-

$$M_X(t) = \frac{\theta}{\theta - t}.$$

$$M_X'(t) = \frac{\theta}{(\theta-t)^2} \quad , \quad M_X''(t) = \frac{2\theta}{(\theta-t)^3}.$$

Now, $E(X) = M_X'(0) = \frac{1}{\theta}$, Mean $= \frac{1}{\theta}$,

$$E(X^2) = M_X''(0) = \frac{2}{\theta^2}.$$

Therefore,

$$Var(X) = E(X^2) - [E(X)]^2$$

$$= \frac{2}{\theta^2} - \frac{1}{\theta^2}$$

$$= \frac{1}{\theta^2}.$$

Now, Variance $(X) = \frac{1}{\theta^2}$.

Note on MGF:- $$M_{X-A}(t) = E\left[e^{t(X-A)}\right] = E\left[\sum_{r=0}^\infty \frac{\{t(X-A)\}^r}{r!}\right]$$

$$= \sum_{r=0}^\infty \frac{t^r}{r!} E(X-A)^r = \sum_{r=0}^\infty \frac{t^r}{r!} \mu_r'(A)$$

$\therefore \mu_r'(A)$ is coefficient of $\frac{t^r}{r!}$ in $M_{X-A}(t)$.

$\therefore \mu_r'$ " " " $\frac{t^r}{r!}$ " $M_X(t)$.

$\therefore \mu_r$ " " " $\frac{t^r}{r!}$ " $M_{X-\mu}(t)$.

2. Normal distribution with parameter μ & σ.

The p.d.f. is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \text{ where } -\infty < x < \infty \text{ \& } \sigma > 0$$

~~Now, $M_X(t)$~~ $= E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx - \frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{2\sigma^2 tx - x^2 + 2x\mu - \mu^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2 - 2\sigma^2 tx)} dx$$

ANS:- considering $z = \frac{x-\mu}{\sigma}$, $x = \sigma z + \mu$

$$M_X(t) = E(e^{tX}) = E(e^{t\sigma z + t\mu})$$

$$= E(e^{t\sigma z + t\mu})$$

$$= e^{t\mu} E(e^{t\sigma z})$$

$$= e^{t\mu} \, M_Z(t\sigma) \quad \text{——} (1)$$

Now, $M_Z(t\sigma) = E(e^{t\sigma z})$

$$= \int_{-\infty}^{\infty} e^{t\sigma z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\sigma z - z^2/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{t^2\sigma^2}{2} - \frac{(z-t\sigma)^2}{2}\right] dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{t^2\sigma^2/2} \int_{-\infty}^{\infty} \exp\left[-\frac{(z-t\sigma)^2}{2}\right] dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{t^2\sigma^2/2} \int_{-\infty}^{\infty} e^{-\frac{m^2}{2}} dm$$

$$= e^{t^2\sigma^2/2} \quad \text{——} (2)$$

Now,
$t\sigma z - z^2/2$
$= \frac{t^2\sigma^2}{2} - \frac{(z-t\sigma)^2}{2}$

Let, $z - t\sigma = m$
$dz = dm$

since $\frac{1}{\sqrt{2\pi}} \cdot e^{-m^2/2}$ is the p.d.f. of a standard normal random variable.

So, $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-m^2/2}\, dm = 1$

Now, Putting ② in eqn. ①,

$$M_X(t) = e^{t\mu} \cdot M_Z(t\sigma)$$
$$= e^{t\mu} \cdot e^{t^2\sigma^2/2}$$
$$= e^{t\mu + \frac{t^2\sigma^2}{2}}$$

Now, differentiating $M_X(t)$, we obtain

$$M_X'(t) = \left(\mu + \frac{2t\sigma^2}{2}\right) \cdot e^{t\mu + \frac{t^2\sigma^2}{2}}$$
$$= (\mu + \sigma^2 t) \cdot e^{t\mu + t^2\sigma^2/2}$$

which upon differentiation gives

$$M_X''(t) = (\mu + \sigma^2 t)^2 \cdot e^{t\mu + t^2\sigma^2/2} + \sigma^2 \cdot \exp\left[t\mu + \tfrac{1}{2}t^2\sigma^2\right]$$

Therefore, $E(X) = M_X'(0) = \mu$

$E(X^2) = M_X''(0) = \mu^2 + \sigma^2$

Thus, $Var(X) = E(X^2) - [E(X)]^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$.

∴ Mean of the normal distn. is μ,
the variance of normal distn. is σ^2.

Alternative:-

$$M(t) = E(e^{tX})$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu + \sigma y)} e^{-y^2/2} dy$$
$$= \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\sigma y - y^2/2} dy$$
$$= \frac{\exp(t\mu)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{-(y - t\sigma)^2/2 + t^2\sigma^2/2\} dy$$
$$= \frac{\exp(t\mu + t^2\sigma^2/2)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-(y - t\sigma)^2/2) dy$$
$$= \exp(t\mu + t^2\sigma^2/2)$$

$y = \frac{x-\mu}{\sigma}$, $dy = \frac{dx}{\sigma}$

3. Beta distribution with parameter m, n.

The p.d.f. is given by

$$f(x) = \frac{1}{B(m,n)} x^{m-1} (1-x)^{n-1} \quad , 0<x<1 \ , m>0, n>0$$

ANS:-

$$M_X(t) = E(e^{tX}) = \int_0^1 e^{tx} \cdot \frac{1}{B(m,n)} \cdot x^{m-1} (1-x)^{n-1} dx$$

$$= \frac{1}{B(m,n)} \int_0^1 e^{tx} x^{m-1} (1-x)^{n-1} dx$$

$$= \frac{1}{B(m,n)} \int_0^1 \left[\sum_k \frac{(tx)^k}{k!} \right] x^{m-1} (1-x)^{n-1} dx$$

$$= \sum_k \frac{t^k}{k!} \int_0^1 x^k f(x) dx$$, since the inner sum is finite

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot \mu_k' \quad \text{——— (1)}$$

Now, kth raw moment,

$$\mu_k' = \frac{1}{B(m,n)} \int_0^1 x^{k+m-1} (1-x)^{n-1} dx$$

$$= \frac{1}{B(m,n)} \cdot B(k+m, n)$$

$$= \frac{\Gamma m+k \, \cancel{\Gamma n} \, \Gamma m+n}{\Gamma m \, \cancel{\Gamma n} \, \Gamma m+k+n} = \frac{\Gamma m+k \, \Gamma m+n}{\Gamma m \, \Gamma m+k+n} \quad \text{— (2)}$$

Putting (2) in equation (1),

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mu_k'$$

$$= \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot \frac{\Gamma m+k \, \Gamma m+n}{\Gamma m \, \Gamma m+n+k}$$

In particular, $\mu_k' = \dfrac{\Gamma m+k \; \Gamma m+n}{\Gamma m \; \Gamma m+n+k}$

for, $k=1$, $\mu_1' = \dfrac{\Gamma m+n \; \Gamma m+1}{\Gamma m \; \Gamma m+n+1} = \dfrac{m}{m+n} = E(X)$

$$\mu_2' = \frac{\Gamma m+2 \; \Gamma m+n}{\Gamma m \; \Gamma m+n+2} = \frac{(m+1)! \; (m+n-1)!}{(m-1)! \; (m+n+1)!}$$

$$= \frac{(m+1)\,(m+n-1)! \times m}{(m+n+1)(m+n)(m+n-1)!}$$

$$= \frac{m(m+1)}{(m+n)(m+n+1)}$$

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

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

$$\cancel{= \frac{1}{(m+n)}\left[\frac{m+1}{m+n+1} - \right]}$$

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

$$= \frac{m}{(m+n)^2(m+n+1)}\left\{(m+1)(m+n) - m(m+n+1)\right\}$$

$$= \frac{mn}{(m+n)^2(m+n+1)}$$

$\therefore$ Mean $= \dfrac{m}{m+n}$, Variance $= \dfrac{mn}{(m+n)^2(m+n+1)}$

NOTE: →

Calculation of moments from MGF : —

$M_{X-A}(t) = E\left[e^{t(X-A)}\right]$; $\dfrac{d}{dt} M_{X-A}(t)\Big|_{t=0} = E(X-A) = \mu_1'(A)$.

$\dfrac{d^2}{dt^2} M_{X-A}(t)\Big|_{t=0} = E(X-A)^2 = \mu_2'(A)$

$\therefore \dfrac{d^r}{dt^r} M_{X-A}(t)\Big|_{t=0} = E(X-A)^r = \mu_r'(A) =$ rth order moment about zero.

if $A=0$, $M_{X-A}^{(r)}(0) = \dfrac{d^r}{dt^r} M_X(t)\Big|_{t=0} = \mu_r' =$ rth order raw moment

if $A=\mu$, $\dfrac{d^r}{dt^r} M_{X-\mu}(t)\Big|_{t=0} = \mu_r =$ rth order central moment.

Theorem: The m.g.f. of the sum $S_n = X_1 + \cdots + X_n$ of n independent r.v.s $X_1, X_2, \ldots, X_n$ is

$$M_{S_n}(t) = M_{X_1}(t) \cdots\cdots M_{X_n}(t),$$

where $M_{X_i}(t)$ is the p.m.f. of X_i, provided all the mgfs exist.

Proof:→

$$M_{S_n}(t) = E\left(e^{t\sum_{i=1}^{n} X_i}\right) = E\left(\prod_{i=1}^{n} e^{tX_i}\right)$$

$$= \prod_{i=1}^{n} E\left(e^{tX_i}\right), \text{ as } X_i\text{'s are independent.}$$

EXAMPLE: 1.1. X is normally distributed with mean μ and s.d. σ.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty \leq x \leq \infty$$

Then m.g.f. of X about mean (μ) is

$$M_{X-\mu}(t) = E\left(e^{t(X-\mu)}\right)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(x-\mu)} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \text{for every } t.$$

$$= \frac{e^{t^2\sigma^2/2}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)-\sigma^2 t]^2} dx$$

$$= e^{t^2\sigma^2/2} = \sum_{r=0}^{\infty} \left[\frac{t^2\sigma^2}{2}\right]^r \Big/ r!$$

Thus, $E(X-\mu) = 0$ or, $E(X) = \mu$.

And $E(X-\mu)^{2r+1} = 0$, $E(X-\mu)^{2r} = \mu_{2r} = \dfrac{\sigma^{2r}(2r)!}{2^r \times r!}$

$$= (2r-1)(2r-3)\ldots 3 \times 1 \times \sigma^{2r}$$

$\times$ for $r = 1, 2, \ldots\ldots$

NOTE:- The relation between the p.g.f. & the m.g.f. of a non-negative integer-valued random variable X is given by

$$M_X(t) = P_X(e^t).$$

If a distribution is specified by its m.g.f., then the moments (about origin) can be obtained as follows:

$$\mu_k' = \left.\frac{d^k M_X(t)}{dt^k}\right]_{t=0}$$

☆ **8.29.** If X has the negative binomial distribution with p.m.f. $f(x) = \binom{x+r-1}{x} p^r q^x$ if $x \geq 0$

$= 0$ o/w

where r is a positive integer, $0 < p < 1$, and $q = 1-p$, S.T. the m.g.f. of X exists. Also obtain the m.g.f.

ANS:-

$$M_X(t) = E(e^{tX})$$
$$= \sum_{x=0}^{\infty} \exp(tx) f(x)$$
$$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} (qe^t)^x$$

is an absolutely convergent series for $qe^t < 1$ i.e. for $t > \ln q$.

Hence the m.g.f. exists. Also,

$$M_X(t) = p^r (1-qe^t)^{-r}, \text{ for } t < -\ln q.$$

Alternative:-

We know,

$$M_X(t) = P_X(e^t)$$

We know,

$$P_X(t) = p^r (1-qt)^{-r}, \text{ for } |qt| < 1$$

Now,

$$M_X(t) = p^r (1-qe^t)^{-r}, \text{ for } t < -\ln q.$$

☆ **8.30.** Suppose X is a non-negative r.v. for which for some positive integer r, $E(X^r)$ does not exist. S.T. $E(e^{tX})$ exists iff $t \leq 0$. Does the m.g.f. exist?

ANS:-

$$e^{tX} = \sum_{r=0}^{\infty} \frac{(tX)^r}{r!}$$

for $t > 0$, then,

$$\exp(tx) > \frac{t^r}{r!} \cdot x^r \quad (\text{for } x > 0)$$

$$\Rightarrow \int_0^{\infty} \exp(tx) dF(x) = \infty, \text{ since } \int_0^{\infty} x^r dF(x) = \infty$$

On the other hand, for $t \leq 0$,

$$\Leftrightarrow \quad 0 \leq \exp(tx) \leq 1$$

$$\Leftrightarrow 0 \leq \int_0^{\infty} \exp(tx) dF(x) \leq 1.$$

As such, $E(e^{tX})$ is defined for $t \leq 0$.

Hence, the p.m.f. does not exist, since we can't get any

$$h > 0 \ni \int_{-\infty}^{\infty} e^{tx} f(x) dx < \infty \text{ for } |t| \leq h$$

☆**8.31.** S.T. the m.g.f. does not exist for the continuous distribution with p.d.f.

$$f(x) = \frac{1}{4x^5} \text{ if } 0 < x < \infty$$
$$= 0 \quad \text{ow.}$$

ANS:→ In this case, $\int_0^\infty |x^5| f(x)\,dx = \frac{1}{4}\int_0^\infty x\,dx$
$$= \infty$$

. Here $E(X^5)$ does not exist.
$\Rightarrow E(e^{tX})$ does not exist for $t > 0$.
Hence, m.g.f. does not exist.

☆**8.32.** S.T. the m.g.f. does not exist for a Cauchy distribution.

ANS:- $f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$, $-\infty < x < \infty$

$$E|X|^r = \frac{2}{\pi}\int_0^\infty x^r \frac{1}{1+x^2}\,dx$$

$$z = \frac{1}{1+x^2}$$

$$= \frac{1}{\pi}\int_0^1 z^{\frac{(1-r)}{2}-1}(1-z)^{[(r+1)/2]-1}\,dz$$

which is a beta function & exists for $r < 1$ and diverges for $r \geq 1$.

$\therefore E|X|^r$ does not exist for $r \geq 1$
$\Rightarrow E(e^{tX})$ does not exist for $r \geq 1$
$\therefore$ m.g.f does not exist.

Alt. method:-

Now, if $t > 0$, $\int_{-\infty}^\infty e^{tx} f(x)\,dx > \int_0^\infty e^{tx} f(x)\,dx > \int_0^\infty tx f(x)\,dx$

$$= \frac{1}{\pi}\int_0^\infty \frac{tx}{1+x^2}\,dx$$

$$= \frac{t}{2\pi}\left[\ln(1+x^2)\right]_0^\infty$$

$$= \infty.$$

for $t > 0$, m.g.f. does not exist.

If $t \leq 0$, $0 \leq \int_{-\infty}^\infty e^{tx} f(x)\,dx \leq 1$.

☆ 8.33. S.T. the m.g.f. of the uniform distribution over the interval (α, β) is

$$M(t) = \frac{e^{\beta t} - e^{\alpha t}}{(\beta - \alpha) t}, \text{ if } t \neq 0$$
$$= 1 \qquad \text{if } t = 0.$$

ANS:→ We know, $\int_{-\infty}^{\infty} e^{tx} f(x)\,dx$ is absolutely convergent for every t. Hence the m.g.f. exists,

Also, $M_X(t) = E(e^{tX})$

$$= \int_{\alpha}^{\beta} e^{tx} \cdot \frac{1}{\beta - \alpha}\,dx$$

$$= \frac{e^{\beta t} - e^{\alpha t}}{(\beta - \alpha) t} \quad \text{if } t \neq 0$$

and if $t = 0$,

$$M_X(0) = \int_{\alpha}^{\beta} \frac{1}{\beta - \alpha}\,dx$$

$$= 1, \text{ for } t = 0.$$

☆ 8.35. For a gamma distn. with p.d.f.

$$f(x) = \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1} \text{ if } x > 0 \ ; \ \alpha > 0, p > 0,$$

obtain the m.g.f. and find the moment about 0 (μ_r'), $r \geq 1$

ANS:— $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} e^{-(\alpha - t)x} x^{p-1}\,dx < \infty$ if $t < \alpha$

Hence the m.g.f. exists.

$$M_X(t) = \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} e^{-(\alpha - t)x} x^{p-1}\,dx$$

$(\alpha - t)x = y$

$\Rightarrow x = \frac{y}{\alpha - t}$

$dx = \frac{dy}{(\alpha - t)}$

$\therefore x^{p-1} = \frac{y^{p-1}}{(\alpha - t)^{p-1}}$

$$= \frac{\alpha^p}{\Gamma(p)} \int_0^{\infty} e^{-y} y^{p-1} \cdot \frac{1}{(\alpha - t)^p} \cdot dy$$

$$= \frac{\alpha^p}{\Gamma(p)(\alpha - t)^p} \int_0^{\infty} e^{-y} y^{p-1}\,dy$$

$$\therefore M_X(t) = \frac{\alpha^p \cdot \cancel{\Gamma(p)}}{\cancel{\Gamma(p)}\,(\alpha - t)^p} = \left(1 - \frac{t}{\alpha}\right)^{-p}, \text{ for } t < \alpha,$$

Since,

$$M_X(t) = 1 + p\frac{t}{\alpha} + \frac{p(p+1)}{2!}\left(\frac{t}{\alpha}\right)^2 + \cdots\cdots + \frac{p(p+1)\cdots(p+r-1)}{r!}\left(\frac{t}{\alpha}\right)^r + \cdots\cdots \infty$$

$\therefore \mu_r' =$ coefficient of $t^r/r!$ in $M_X(t)$

$$= \frac{p(p+1)\cdots(p+r-1)}{\alpha^r}.$$

8.34. If x has the Laplace distribution with p.d.f.

$$f(x) = \frac{\alpha}{2} \exp(-\alpha|x|), \quad -\infty < x < \infty, \quad \alpha > 0,$$

find the m.g.f. of X. For what value of t is it defined. Hence obtain $E(X)$ and $Var(X)$.

ANS:-

$$E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$$

$$= \int_{-\infty}^{\infty} e^{tx} \cdot \frac{\alpha}{2} e^{-\alpha|x|} dx$$

$$= \frac{\alpha}{2}\left[\int_{-\infty}^{0} e^{x(t+\alpha)} dx + \int_{0}^{\infty} e^{x(t-\alpha)} dx\right]$$

the integrals are convergent for $|t|<\alpha$; , defined for $|t| < \alpha$

As such $E(e^{tX})$ exists for $|t|<\alpha$, impling that the m.g.f. exists.

$$\therefore M_X(t) = E(e^{tX}) = \frac{\alpha}{2}\left\{\left[\frac{1}{t+\alpha} + \frac{1}{\alpha - t}\right]\right\}$$

$$= \frac{1}{2}\left[\left(1+\frac{t}{\alpha}\right)^{-1} + \left(1-\frac{t}{\alpha}\right)^{-1}\right]$$

$$= \frac{1}{2}\left[\left(1 - \frac{t}{\alpha} + \frac{t^2}{\alpha^2} - \cdots \infty\right) + \left(1 + \frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \frac{t^3}{\alpha^3} + \cdots\right)\right]$$

$$= \left[1 + \frac{t^2}{\alpha^2} + \frac{t^4}{\alpha^4} + \frac{t^6}{\alpha^6} + \cdots \infty\right]$$

$$\therefore M_X(t) = \frac{1}{1 - \frac{t^2}{\alpha^2}} = \frac{\alpha^2}{\alpha^2 - t^2}.$$

$$\therefore M_X'(t) = \frac{2t\alpha^2}{(\alpha^2 - t^2)^2}$$

$$\therefore M_X''(t) = \frac{2\alpha^2(\alpha^2-t^2)^2 + 32\alpha^4 t^5}{(\alpha^2 - t^2)^4}$$

$$\therefore E(X) = M_X'(0) = 0$$

$$\therefore E(X^2) = M_X''(0) = \frac{2\alpha^{\cancel{6}}}{\alpha^{\cancel{8}2}} = \frac{2}{\alpha^2}$$

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

$$= \frac{2}{\alpha^2}.$$

Some more examples: →

☆1. From the p.g.f. $P_X(t)$ of $X = 0, 1, 2, \ldots$, find the p.g.f. of $Y = 3X + 2$.

Ans: →
$$P_Y(t) = E(t^Y) = E\left(t^{3X+2}\right) = E(t^{3X} \cdot t^2)$$
$$= t^2 E(t^{3X})$$
$$= t^2 P_X(t^3).$$

☆2. Using the p.g.f. prove that the r.v. S_N of $S_N = X_1 + X_2 + \cdots + X_N$ has the variance given by $Var(S_N) = E(N)\, Var(X) + Var(N) E^2(X)$ [Wald's eqn. on random sum]

Soln.
$$Var(S_N) = P_{S_N}''(1) + P_{S_N}'(1) - P_{S_N}'(1)^2, \quad \text{[Feller]} \quad \text{------ (1)}$$

where,
$$P_{S_N}(t) = P_N(P_X(t))$$
$$P_{S_N}'(t) = P_X'(t)\, P_N'(P_X(t))$$
$$P_{S_N}''(t) = P_X''(t) P_N'(P_X(t)) + P_X'(t)^2 P_N''(P_X(t)).$$
(differentiating w.r.t. t)

Hence, we have,
$$P_{S_N}'(t) = E(S_N) = P_X'(1) P_N'(P_X(1)) \quad \left[\because P_X(1) = \sum_{x \in S} f(x) = 1\right]$$
$$= P_X'(1) P_N'(1) = E(X) E(N),$$

$$P_{S_N}''(1) = P_X''(1) P_N'(1) + [P_X'(1)]^2 P_N''(1)$$
$$= E[X(X-1)] E(N) + E(N^2) E[X(X-1)] \quad \text{----- (2)}$$

Putting (2) in (1), we have the result.

8.18. Let X and Y (two r.v.s) assume non-negative integral values and

$$P[X=n, Y=k] = \frac{e^{-\lambda}\lambda^n}{n!}\binom{n}{k}p^k q^{n-k} \text{ if } n \geq 0 \text{ and } k=0,1,\dots,n$$

Find the p.g.f. of X, Y.

ANS: → The joint p.g.f. is

$$P(t_1,t_2) = E\left(t_1^X t_2^Y\right)$$
$$= E\left[t_1^X E\left(t_2^X \mid X\right)\right]$$
$$= E\left[t_1^X (q+pt_2)^X\right],$$ since for, $X=x$, the conditional distn. of Y is $Bin(x,p)$.

$$= e^{-\lambda(1-t_1q-t_1t_2p)}$$

⇒ p.g.f.s of X and Y are, $P_X(t_1) = P(t_1,1) = e^{\lambda(t_1-1)}$

& $P_Y(t_2) = P(1,t_2) = e^{\lambda p(t_2-1)}$.

☆ 3) Find the p.g.f. and hence mean & variance of the following distributions: (a) binomial, (b) Poisson, (c) geometric, (d) Pascal. Moreover deduce the corresponding m.g.f and characteristic functions.

Soln. →

[NOTE: If $P(t)$: Probability Generating function, $M(t)$: Moment Generating function, $\phi(t)$: characteristic function; exists for a variable X, then the following relations hold:

$$M_X(t) = P_X(e^t) = \phi_X(-it)$$
$$\phi_X(t) = M_X(it) = P_X(e^{it})$$]

& $E(X^r) = M^{(r)}(0)$, where $M^{(r)}(t)$ denotes the rth derivative of M. Furthermore, it determines uniquely (characterizes) the distribution of X.

(a) BINOMIAL DISTRIBUTION:—

$$f_X(x \mid n,p) = \binom{n}{x}p^x q^{n-x} \quad ; x \geq 0,\ p+q=1,\ p,q>0$$

$$P_X(t) = E(t^X) = \sum_{x=0}^{n} t^x \cdot \binom{n}{x}p^x q^{n-x}$$
$$= \sum_{x=0}^{n}\binom{n}{x}(tp)^x q^{n-x}$$
$$= (q+pt)^n$$

$\therefore M(t) = P_X(e^t) = (q+pe^t)^n$; $\phi(t) = M(it) = (q+pe^{it})^n$,

Mean $(\mu) = P_X'(1) = M_X'(0) = i^{-1}\phi'(0)$

$\therefore \frac{d}{dt}[P_X(t)] = n(q+pt)^{n-1}\cdot p = P_X'(t)$

$\therefore P_X'(1) = n(p+q)^{n-1}\cdot p$

$= pn$, as $p+q=1$.

$\therefore P_X''(t) = np\cdot(n-1)(q+pt)^{n-2}\cdot p = np^2(n-1)(q+pt)^{n-2}$

$\therefore P_X''(1) = n(n-1)p^2$; $Var(X) = E(X^2) - [E(X)]^2 = P_X''(1) + P_X'(1) - \{P_X'(1)\}^2$

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

$= np(1-p)$

$= npq$.

$\mu = np$, $\sigma^2 = pqn$.

(b) POISSON DISTRIBUTION :-

$$f_X(x|\lambda) = e^{-\lambda}\cdot\frac{\lambda^x}{x!}, \quad x \geq 0, \ \lambda > 0.$$

$$P(t) = E(t^X) = \sum_{x=0}^{\infty} t^x \cdot e^{-\lambda}\cdot\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\cdot e^{\lambda t}$$

$$= e^{\lambda(t-1)}$$

$$M(t) = P(e^t) = e^{\lambda(e^t-1)}.$$

$$\phi(t) = P(e^{it}) = e^{\lambda(e^{it}-1)}$$

$\therefore M_X'(t) = \lambda e^{\cdots t}$ $\therefore M_X'(0) = \lambda$;

$\therefore M_X''(t) = \frac{1}{dt}\{e^{\lambda(e^t-1)}\cdot\lambda e^t\}$; $\lambda + \lambda^2 = M_X''(0)$

$\therefore Var(X) = M_X''(0) - \{M_X'(0)\}^2 = \lambda$

$\mu = \sigma^2 = \lambda$.

(c) GEOMETRIC DISTRIBUTION :-

$$f_X(x|p) = pq^{x-1}, \quad x \geq 0, \ p+q=1, \ p,q>0$$

$$P(t) = E(t^X) = \sum_{x=1}^{\infty} t^x\cdot pq^{x-1}$$

$$= pt\sum_{x=1}^{\infty}(qt)^{x-1}$$

$$= pt\left(\frac{1}{1-qt}\right) = \frac{pt}{1-qt}$$

$$M_X(t) = P(e^t) = \frac{pe^t}{1-qe^t}, \quad \phi(t) = \frac{pe^{it}}{1-qe^{it}}$$

$$\therefore P_X'(t) = \frac{p(1-qt)+q\cdot pt}{(1-qt)^2} = \frac{p}{(1-qt)^2}$$

$\therefore P_X'(1) = \frac{p}{p^2} = \frac{1}{p} = E(X)$

$\therefore P_X''(t) = p\{-2(1-qt)^{-3}\cdot -q\} = 2pq(1-qt)^{-3} = \frac{2pq}{(1-qt)^3}$

$\therefore P_X''(1) = \frac{2pq}{(1-q)^3} = \frac{2q}{p^2}$

$\therefore Var(X) = P_X''(1) + P_X'(1) - \{P_X'(1)\}^2 = \frac{2q}{p^2} - \frac{1}{p^2} + \frac{1}{p} = \frac{2q+p-1}{p^2}$

$= q/p^2$

$\therefore \mu = \frac{1}{p}$, $\sigma^2 = \frac{q}{p^2}$.

Scanned by CamScanner

PASCAL DISTRIBUTION :-

$$f_X(x|r,p) = \binom{r+x-1}{x} p^r q^x \quad , x \geq 0, \quad p, q > 0, p+q=1$$

$$P(t) = E(t^X)$$
$$= \sum_{x=0}^{\infty} t^x \binom{x+r-1}{x} p^r q^x$$
$$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} (qt)^x$$
$$= p^r (1-qt)^{-r} = \left(\frac{p}{1-qt}\right)^r$$

$$M(t) = \left(\frac{p}{1-qe^t}\right)^r \quad ; \quad \phi(t) = \left(\frac{p}{1-qe^{it}}\right)^r$$

$$M_X'(t) = p^r(-r)(1-qe^t)^{-r-1}.(-qe^t)$$

$$M_X'(0) = p^r p^{-r-1}.rq = \frac{rq}{p}$$

$$M_X''(t) = p^r.r\{qe^t.\{-(r+1)\}(1-qe^t)^{-r-2}.(-qe^t) + (1-qe^t)^{-r-1}.(qe^t)\}$$

$$M_X''(0) = p^r.r\{q\{-(r+1)\}p^{-r-2}.(-q) + p^{-r-1}.(+q)\} = \frac{(r^2+r)q^2}{p^2} + \frac{rq}{p}$$

$$\therefore Var(X) = M_X''(0) - [M_X'(0)]^2 = \cancel{\frac{r^2q^2}{p^2}} + \frac{rq^2}{p^2} + \frac{rq}{p} - \cancel{\frac{r^2q^2}{p^2}}$$

$$= \frac{rq^2 + rpq}{p^2} = \frac{rq(q+p)}{p^2} = \frac{rq}{p^2}.$$

$$\therefore E(X) = \frac{rq}{p}, \quad \sigma^2 = \frac{rq}{p^2}.$$

SUMS OF INDEPENDENT RANDOM VARIABLES :-

Theorem: Let $X_1, \dots, X_n$ be indep. r.v.s with m.g.f. $M_{X_1}(t), M_{X_2}(t), \dots, M_{X_n}(t)$: the m.g.f. of $X_1 + \dots + X_n$ is given by

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

Proof:-

Let $W = X_1 + X_2 + \dots + X_n$; then by definition

$$M_W(t) = E(e^{tW}) = E(e^{tX_1 + tX_2 + \dots + tX_n})$$
$$= E(e^{tX_1}.e^{tX_2}.\dots.e^{tX_n})$$
$$= E(e^{tX_1}) \cdot E(e^{tX_2}) \dots E(e^{tX_n})$$
$$= M_{X_1}(t) \; M_{X_2}(t) \dots M_{X_n}(t).$$

as $X_1, \dots, X_n$'s are independent.

EXAMPLE :1. $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$ } independent, **Poisson Distn.**

$$M_X(t) = e^{\lambda_1(e^t-1)}$$

$$M_Y(t) = e^{\lambda_2(e^t-1)}$$

$$\therefore M_{X+Y}(t) = M_X(t)M_Y(t)$$
$$= e^{(\lambda_1+\lambda_2)(e^t-1)}$$

which is the m.g.f. of $P(\lambda_1+\lambda_2)$.

$$\therefore X+Y \sim P(\lambda_1+\lambda_2)$$

2. $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ **Normal Distn.**

$$M_X(t) = e^{t\mu_1 + \frac{t^2\sigma_1^2}{2}}$$

$$M_Y(t) = e^{t\mu_2 + \frac{t^2\sigma_2^2}{2}}$$

$$\therefore M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$
$$= e^{t(\mu_1+\mu_2) + \frac{t^2}{2}(\sigma_1^2+\sigma_2^2)}$$

$$\therefore X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2).$$

NOTE:

3. Sums of independent geometric random variables are negative binomial.
4. Sums of independent negative binomial random variables are negative binomial.
5. Sums of independent exponential random variables are gamma.
6. Sums of independent gamma random variables are gamma.

B.37 An r.v. X has the Poisson distribution with parameter λ while the conditional distribution of Y for given $X = x$ ($x = 0, 1, 2, \ldots$) is Poisson with parameter μx. Obtain the p.g.f of the (unconditional) distribution of Y and hence its mean & variance.

ANS:— $X \sim \text{Poisson}(\lambda)$

$Y | X = x \sim \text{Poisson}(\mu x)$

To find p.g.f. of the unconditional distn. of Y.

$$P_X(t) = E(t^X) = e^{\lambda(t-1)}$$

$$P_{Y|X}(t) = e^{\mu x(t-1)} = E(t^Y | X = x)$$

$$P_Y(t) = E(t^Y) = \cancel{E_X(E(t^Y|X))}\ E[E(t^Y|X)]$$

$$= E\left(e^{\mu x(t-1)}\right)$$

$$= E\left[e^{\mu(t-1)\cdot X}\right]$$

$$= e^{\lambda(e^{\mu(t-1)}-1)}$$

Mean, variance:—

$$P_Y'(t) = \lambda\mu \cdot e^{\lambda(e^{\mu(t-1)}-1)} \cdot e^{\mu(t-1)}$$

$$P_Y'(1) = \lambda\mu = \text{Mean}.$$

$$e^{\mu(t-1)} = z$$

Cumulant Generating function:—

Suppose for a r.v. X, the m.g.f. $M_X(t)$ is defined then $\ln M_X(t)$ is also defined and is called the c.g.f. of X.
It is denoted by $K_X(t)$.
If the c.g.f. is expanded as a power series in t then the co-efficient of $\frac{t^r}{r!}$ in that series is called the rth cumulant of X, it is denoted by κ_r.

Characteristic function:—

We have already noted that the m.g.f. of a r.v. does not always exist, in such cases also the c.f. exists.

Suppose $Z = X + iY$, $i = \sqrt{-1}$

then we say that, $E(Z) = E(X) + iE(Y)$

If X is an ordinary r.v. then $E(e^{itX})$ is called the characteristic function of X and is denoted by $\phi_X(t)$.

$$\int_{-\infty}^{\infty} |e^{itx}| dF(x) < \infty, \text{ since}$$

$$|e^{itx}| = |\cos x + i\sin x|$$
$$= \sqrt{\sin^2 x + \cos^2 x}$$
$$= 1,$$

$$\therefore \int_{-\infty}^{\infty} |e^{itx}| dF(x) = \int_{-\infty}^{\infty} dF(x) = 1, \text{ i.e. finite.}$$

$\therefore \phi_X(t) = E(e^{itX})$ always exist.

- **Alt. Definition of CF:** ⟶

Let X be an RV. The complex valued function ϕ defined on $\mathbb{R}$ by

$$\phi(t) = E(e^{itX}) = E(\cos tX) + iE(\sin tX), \; t \in \mathbb{R}$$

where $i = \sqrt{-1}$ is the imaginary unit, is called the characteristic function (CF) of RV X.

DISCRETE DISTRIBUTION

Distribution	p.m.f.	Mean & variance	M.g.f.	Special Note
Bernoulli (p)	$P(X=x\|p)$ $=p^x(1-p)^{1-x}$ $x=0,1$ $0<p<1$	$E(X)=p$ $Var(X)=p(1-p)$	$M_X(t)$ $=pe^t+(1-p)$	—
Binomial (n,p)	$P(X=x\|p)$ $=\binom{n}{x}p^x(1-p)^{n-x}$ $x=0,1,2,\dots,n$ $0<p<1,\ q=1-p$	$E(X)=np$ $Var(X)=np(1-p)$	$M_X(t)$ $=(pe^t+q)^n$	can be generalized to multi normal distribution.
Discrete uniform (N)	$P(X=x\|N)$ $=\frac{1}{N}$ $x=1,2,\dots,N$ $N=1,2,\dots$	$E(X)=\frac{N+1}{2}$ $V(X)=\frac{N^2-1}{12}$	$M_X(t)$ $=\frac{1}{N}\sum_{i=1}^{N}e^{it}$	—
Geometric (p)	$P(X=x\|p)$ $=p(1-p)^{x-1}$ $x=1,2,\dots$ $0\le p\le 1$	$E(X)=\frac{1}{p}$ $Var(X)=\frac{1-p}{p^2}$	$M_X(t)=\frac{pe^t}{1-qe^t}$ defined for $t<-\ln q$	$Y=X-1$ is negative binomial$(1,p)$. The distn. has memoryless property, $P(X>s\mid X>t)$ $=P(X>s-t)$
Hypergeometric (N,M,n)	$P(X=x\|N,M,n)$ $=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ $x=\max\{0,n-N+m\}$ (1) $\min\{N,n\}$ $M,N,n\ge 0$	$E(X)=\frac{nM}{N}$ $V(X)=\frac{nM}{N}\cdot\frac{(N-M)(N-n)}{N(N-1)}$		If $k<<M$ & $k<<N$, then the range $x=0,1,2,\dots,k$ will be appropriate. Alt. form of the p.m.f. Let $M=np$, $0<p<1$ $\therefore N-M=Nq$ $\therefore P(X=x\|N,p,n)$ $=\frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$
Negative Binomial (r,p)	$P(X=x\|r,p)$ $=\binom{r+x-1}{x}p^r(1-p)^x$ $x=0,1,2,\dots$ $0\le p\le 1$	$E(X)=\frac{rq}{p}$ $V(X)=\frac{rq}{p^2}$	$M_X(t)$ $=p^r(1-qe^t)^{-r}$ defined for $t<-\ln q$	Alt. form $P(Y=y\|r,p)$ $=\binom{y-1}{r-1}p^rq^{y-r}$ $y=r,r+1,\dots$ where $Y=X+r$
Poisson (λ)	$P(X=x\|\lambda)$ $=e^{-\lambda}\cdot\frac{\lambda^x}{x!}$ $x=0,1,\dots$ $\lambda\ge 0$	$E(X)=\lambda$ $V(X)=\lambda$	$M_X(t)$ $=e^{\lambda(e^t-1)}$	

CONTINUOUS DISTRIBUTION

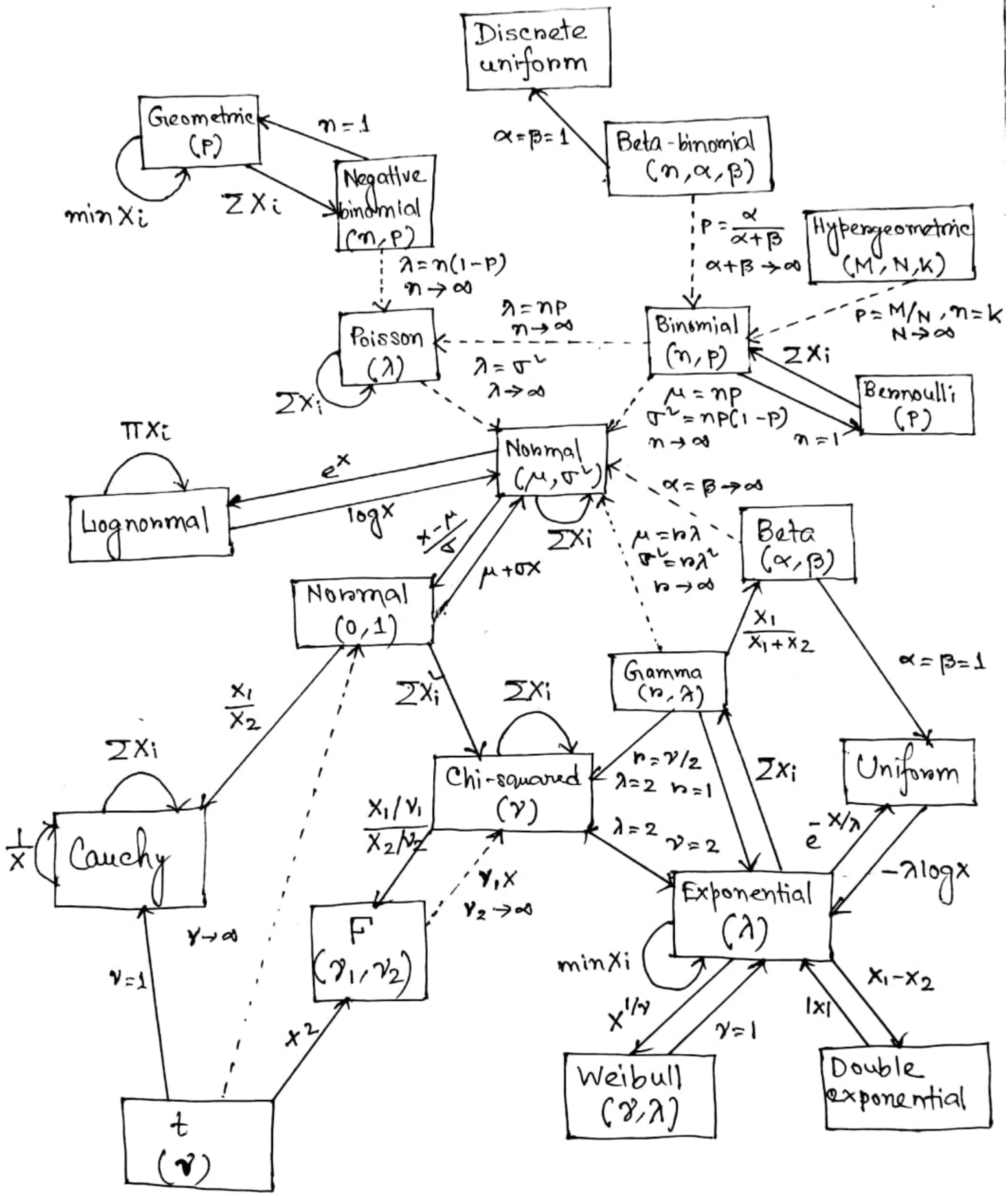
Name.	p.d.f.	Expectation & variance	Mgf	Special notes
Beta (α,β)	$f(x\|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} \cdot x^{\alpha-1}(1-x)^{\beta-1}$ $0\le x\le 1$, $\alpha>0$, $\beta>0$	$E(X)=\frac{\alpha}{\alpha+\beta}$ $V(X)=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$M_X(t) = 1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$	$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
Cauchy (θ,σ)	$f(x\|\theta,\sigma) = \frac{\sigma}{\pi}\cdot\frac{1}{\sigma^2+(x-\theta)^2}$ $-\infty<x<\infty$ $-\infty<\theta<\infty$, $\sigma>0$	don't exist	doesn't exist	Special case of t-distribution when degree of freedom $n=1$.
Chi-square	$f(x\|n) = \frac{1}{\Gamma(\frac{n}{2})2^{n/2}} \cdot e^{-x/2}x^{\frac{n}{2}-1}$ $0\le x<\infty$, $n=1,2,\ldots$	$E(X)=n$ $V(X)=2n$	$M_X(t)=(1-2t)^{-n/2}$ defined for $t<\frac{1}{2}$	Special case of gamma distribution
Exponential (θ)	$f(x\|\theta)=\theta e^{-x\theta}$ $\theta>0$, $0\le x<\infty$	$E(X)=\frac{1}{\theta}$ $V(X)=\frac{1}{\theta^2}$	$M_X(t)=(1-\beta t)^{-1}$ defined for $t<\frac{1}{\beta}$	Special case of Gamma distn. Alt. form $f(x\|\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}}$ $0\le x<\infty$, $\theta>0$
Double exponential (μ,σ)	$f(x\|\mu,\sigma) = \frac{1}{2\sigma}\cdot e^{-\|x-\mu\|/\sigma}$ $-\infty<x<\infty$ $-\infty<\mu<\infty$, $0<\sigma<\infty$	$E(X)=\mu$ $V(X)=2\sigma^2$	$M_X(t)=\frac{e^{\mu t}}{1-\sigma^2t^2}$ Defined for $\|t\|<\frac{1}{\sigma}$	Also known as Laplace Distribution
Gamma (α,β)	$f(x\|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\cdot x^{\alpha-1}e^{-\frac{x}{\beta}}$ $0\le x<\infty$, $\alpha,\beta>0$	$E(X)=\alpha\beta$ $V(X)=\alpha\beta^2$	$M_X(t)=(1-\beta t)^{-\alpha}$ $t<\frac{1}{\beta}$	When $\alpha=1$, it is exponential distn. If $\alpha=\frac{n}{2}$, $\beta=2$, it is χ^2-distn.
Normal (μ,σ^2)	$f(x\|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $-\infty<x<\infty$, $-\infty<\mu<\infty$, $\sigma>0$	$E(X)=\mu$ $V(X)=\sigma^2$	$M_X(t) = e^{\mu t+\frac{1}{2}\sigma^2t^2}$	Also known as Gaussian distribution.
log-normal (μ,σ^2)	$f(x\|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\}$ $0\le x<\infty$, $-\infty<\mu<\infty$ $0<\sigma<\infty$	$E(X)=e^{\mu+\sigma^2/2}$ $V(X)=e^{2(\mu+\sigma^2)}-e^{2\mu+\sigma^2}$	Does not exist but $EX^n=e^{n\mu+\frac{n^2\sigma^2}{2}}$	—
Uniform (α,β)	$f(x\|\alpha,\beta)=\frac{1}{\beta-\alpha}$, $\alpha\le x\le\beta$	$E(X)=\frac{\beta+\alpha}{2}$ $V(X)=\frac{(\beta-\alpha)^2}{12}$	$M_X(t) = \frac{e^{\beta t}-e^{\alpha t}}{(\beta-\alpha)t}$	$\alpha=0$ & $\beta=1$, then it is a special case of Beta distn. (with parameter 1 & 1)

CONTINUOUS DISTRIBUTION

Name	p.d.f.	Expectation & variance	M.g.f.	Special Notes
Weibull (γ, β)	$f(x\mid\gamma,\beta) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-\frac{x^\gamma}{\beta}}$ $0 \le x < \infty,\ \gamma > 0,\ \beta > 0$	$E(X) = \beta^{1/\gamma}\Gamma(1+\frac{1}{\gamma})$ $V(X) = \beta^{2/\gamma}[\Gamma(1+\frac{2}{\gamma}) - \Gamma^2(1+\frac{1}{\gamma})]$	Exists only for $\gamma \ge 1$, Form is not useful	special case, $\gamma = 1$, exponential
Pareto (α, β)	$f(x\mid\alpha,\beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}$ $\alpha < x < \infty$, $\alpha > 0,\ \beta > 0$	$E(X) = \frac{\beta\alpha}{\beta-1}$ $(\beta > 1)$ $V(X) = \frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)}$ $(\beta > 2)$	Does not exist	—
Logistic (μ, β)	$f(x\mid\mu,\beta)$ $= \frac{1}{\beta} \cdot \frac{e^{-\frac{(x-\mu)}{\beta}}}{[1+e^{-\frac{(x-\mu)}{\sigma}}]^2}$ $-\infty < x < \infty$ $-\infty < \mu < \infty,\ \beta > 0$	$E(X) = \mu$ $V(X) = \frac{\pi^2\beta^2}{3}$	$M_X(t)$ $= e^{\mu t}(1-\beta t)\Gamma(1+\beta t)$ defined for $\lvert t\rvert < \frac{1}{\beta}$	The d.f. is $F(x) = \frac{1}{1+e^{-\frac{(x-\mu)}{\beta}}}$
$t(\nu)$	$f(x\mid\nu)$ $= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{\nu\pi}} \cdot \frac{1}{[1+(\frac{x^2}{\nu})]^{\frac{\nu+1}{2}}}$ $-\infty < x < \infty$ $\nu = 1, 2, \ldots$	$E(X) = 0,\ \nu > 1$ $V(X) = \frac{\nu}{\nu-2},\ \nu > 2$	Does not exist but $EX^n = 0$ if $n < \nu$ and odd $EX^n = \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{\nu-n}{2})}{\sqrt{\pi}\,\Gamma(\frac{\nu}{2})}$ if $n < \nu$ & even	—
$F(\nu_1, \nu_2)$	$f(x\mid\nu_1,\nu_2)$ $= \frac{\Gamma(\frac{\nu_1+\nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \frac{x^{(\frac{\nu_1}{2}-1)}}{[1+(\frac{\nu_1}{\nu_2})x]^{\frac{\nu_1+\nu_2}{2}}}$ $0 \le x < \infty,\ \nu_1, \nu_2 = 1, \ldots$	$E(X) = \frac{\nu_2}{\nu_2-2},\ \nu_2 > 2$ $V(X) = 2\left(\frac{\nu_2}{\nu_2-2}\right)^2 \frac{(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}$, $\nu_2 > 4$	Doesn't exist	—

Scanned by CamScanner

TABLE OF COMMON DISTRIBUTIONS



Relationships among common distributions. Solid lines represent transformations & special cases, dashed lines represent limits.

Bernoulli Distribution

Parameters :	$0 < p < 1 \, , \, p \in \mathbb{R}$
Support :	$k = \{0, 1\}$
pmf :	$q = (1-p)$ for $k = 0$ p for $k = 1$ $\{ p^{k} (1-p)^{1-k} \}$
cdf :	0 for $k < 0$ q for $0 \leq k < 1$ 1 for $k \geq 1$
mean :	p
median:	N/A
mode :	0 if $q > p$ $0, 1$ if $q = p$ 1 if $q < p$
variance:	pq
Skewness :	$\dfrac{q-p}{\sqrt{pq}}$
ex. kurtosis :	$\dfrac{6p^2 - 6p + 1}{p(1-p)}$
entropy :	$-q \ln(q) - p \ln(p)$
mgf :	$q + pe^{t}$
cf :	$q + pe^{it}$
pgf :	$G(z) = q + pz$

BINOMIAL DISTRIBUTION

notation:	$B(n,p)$
parameters:	$n \in \mathbb{N}_0$ – number of trials $p \in [0,1]$ – success probability in each trial
support:	$k \in \{0, \dots, n\}$
pmf:	$\binom{n}{k} p^k (1-p)^{n-k}$
cdf:	$I_{1-p}(n-k, 1+k)$
mean:	np mean > variance
median:	$\lfloor np \rfloor$ or $\lceil np \rceil$
mode:	$\lfloor (n+1)p \rfloor$ or $\lfloor (n+1)p \rfloor - 1$
variance:	$np(1-p)$
Skewness:	$\dfrac{1-2p}{\sqrt{np(1-p)}}$
ex. kurtosis:	$\dfrac{1-6p(1-p)}{np(1-p)}$
entropy:	$\frac{1}{2}\log_2\left(2\pi e np(1-p)\right) + O\left(\frac{1}{n}\right)$
mgf:	$(1-p+pe^t)^n$
cf:	$(1-p+pe^{it})^n$
pgf:	$G(z) = [(1-p)+pz]^n$.

DISCRETE UNIFORM DISTRIBUTION

parameters :	$a \in (\ldots, -2, -1, 0, 1, 2, \ldots)$ $b \in (\ldots, -2, -1, 0, 1, 2, \ldots)$, $b \geq a$ $n = b - a + 1$
support :	$k \in \{a, a+1, \ldots, b-1, b\}$
pmf :	$\frac{1}{n}$ for $a \leq k \leq b$ 0 ow
cdf :	0 for $k < a$ $\frac{\lfloor k \rfloor - a + 1}{n}$ for $a \leq k \leq b$ 1 for $k > b$
mean :	$\frac{a+b}{2}$
median :	$\frac{a+b}{2}$
mode :	N/A
variance :	$\frac{(b-a+1)^2 - 1}{12} = \frac{n^2-1}{12}$,
skewness :	0
ex. kurtosis :	$-\frac{6(n^2+1)}{5(n^2-1)}$
entropy :	$\ln(n)$
mgf :	$\frac{e^{at} - e^{(b+1)t}}{n(1-e^t)}$
cf :	$\frac{e^{iat} - e^{i(b+1)t}}{n(1-e^{it})}$

GEOMETRIC DISTRIBUTION

	No. of trials	no. of failures
parameters	$0 < p \leq 1$ success probability (real)	$0 < p \leq 1$ success probability (real)
Support :	$k \in \{1, 2, 3, \ldots\}$	$k \in \{0, 1, 2, 3, \ldots\}$
pmf :	$(1-p)^{k-1} p$	$(1-p)^k p$
cdf :	$1 - (1-p)^k$	$1 - (1-p)^{k+1}$
Mean :	$\frac{1}{p}$	$\frac{1-p}{p}$
Median :	$\left\lceil \frac{-\log(2)}{\log(1-p)} \right\rceil$ (not unique if $-\log(2)/\log(1-p)$ is an integer)	(variance > mean)
Mode :	1	0
Variance :	$\frac{1-p}{p^2}$	$\frac{1-p}{p^2}$
Skewness :	$\frac{2-p}{\sqrt{1-p}}$	$\frac{2-p}{\sqrt{1-p}}$
Excess Kurtosis :	$6 + \frac{p^2}{1-p}$	$6 + \frac{p^2}{1-p}$
Entropy :	$\frac{-(1-p)\log_2(1-p) - p\log_2 p}{p}$	$\frac{-(1-p)\log_2(1-p) - p\log_2 p}{p}$
mgf :	$\frac{pe^t}{1-(1-p)e^t}$ for $t < -\ln(1-p)$	$\frac{p}{1-(1-p)e^t}$
c.f :	$\frac{pe^{it}}{1-(1-p)e^{it}}$	$\frac{p}{1-(1-p)e^{it}}$

Scanned by CamScanner

NEGATIVE BINOMIAL DISTRIBUTION

notation:	$NB(r, p)$
parameters:	$r>0$ — number of failures until the experiment is stopped (integer, but the definition can also be extended to reals) $p \in (0,1)$ - success probability in each experiment (real).
support:	$k \in \{0,1,2,3,\dots\}$
pmf:	$\binom{k+r-1}{k} \cdot (1-p)^r p^k$, involving a binomial coefficient
cdf:	$1 - I_p(k+1, r)$, the regularized incomplete beta function
mean:	$\frac{pr}{1-p}$ mean < variance.
median:	
mode:	$\begin{cases} \left\lfloor \frac{p(r-1)}{1-p} \right\rfloor & \text{if } r>1 \\ 0 & \text{if } r \le 1 \end{cases}$
variance:	$\frac{pr}{(1-p)^2}$
skewness:	$\frac{1+p}{\sqrt{pr}}$
ex. kurtosis:	$\frac{6}{r} + \frac{(1-p)^2}{pr}$
entropy:	
mgf:	$\left(\frac{1-p}{1-pe^t}\right)^r$ for $t < -\log p$
cf:	$\left(\frac{1-p}{1-pe^{it}}\right)^r$ with $t \in \mathbb{R}$
pgf:	$\left(\frac{1-p}{1-pz}\right)^r$ for $\|z\| < \frac{1}{p}$.

HYPERGEOMETRIC DISTRIBUTION

parameters	$N \in \{1, 2, \ldots\}$ $m \in \{0, 1, 2, \ldots, N\}$ $n \in \{1, 2, \ldots, N\}$
support:	$k \in \{\max(0, n+m-N), \ldots, \min(m, n)\}$
pmf:	$\dfrac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$
cdf:	
mean:	$\dfrac{nm}{N}$
median:	
mode:	$\left\lfloor \dfrac{(n+1)(m+1)}{N+2} \right\rfloor$
variance:	$\dfrac{nm\,(N-n)(N-m)}{N^2\,(N-1)}$
skewness:	$\dfrac{(N-2m)(N-1)^{1/2}(N-2n)}{[nm\,(N-m)(N-n)]^{1/2}\,(N-2)}$
ex. kurtosis:	$\left[\dfrac{N^2(N-1)}{n(N-2)(N-3)(N-n)}\right] \cdot \left[\dfrac{N(N+1)-6N(N-n)}{m(N-m)} + \dfrac{3n(N-n)(N+6)}{N^2} - 6\right]$
entropy:	
mgf:	$\dfrac{\binom{N-m}{n}\,{}_2F_1(-n, -m; N-m-n+1; e^t)}{\binom{N}{n}}$
cf:	$\dfrac{\binom{N-m}{n}\,{}_2F_1(-n, -m; N-m-n+1; e^{it})}{\binom{N}{n}}$

POISSON DISTRIBUTION

notation:	$\mathrm{Pois}(\lambda)$
parameters:	$\lambda > 0$ (real)
support:	$k \in \{0, 1, 2, 3, \ldots\}$
pmf:	$\frac{\lambda^k}{k!} \cdot e^{-\lambda}$
cdf:	$\frac{\Gamma(\lfloor k+1 \rfloor, \lambda)}{\lfloor k \rfloor !}$ for $k \geqslant 0$ or $e^{-\lambda} \sum_{i=0}^{k} \frac{\lambda^i}{i!}$ (where $\Gamma(x, y)$ is the incomplete gamma function and $\lfloor k \rfloor$ is the floor function)
mean:	λ Mean = Variance
median:	$\approx \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$
mode:	$\lceil \lambda \rceil - 1$
variance:	λ
skewness:	$\lambda^{-1/2}$
ex. kurtosis:	λ^{-1}
entropy:	$\lambda[1 - \log(\lambda)] + e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k \log(k!)}{k!}$ (for large λ) $\frac{1}{2}\log(2\pi e \lambda) - \frac{1}{12\lambda} - \frac{1}{24\lambda^2} - \frac{19}{360\lambda^3} + O\left(\frac{1}{\lambda^4}\right)$
mgf:	$\exp(\lambda(e^t - 1))$
cf:	$\exp(\lambda(e^{it} - 1))$

	Beta Distribution	Exponential Distribution
parameters:	$\alpha>0$ shape (real) $\beta>0$ shape (real)	$\lambda>0$ rate, or inverse scale
support:	$x \in (0;1)$	$x \in [0,\infty)$
pdf:	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$	$\lambda e^{-\lambda x}$
cdf:	$I_x(\alpha,\beta)$	$1-e^{-\lambda x}$
mean:	$\frac{\alpha}{\alpha+\beta}$	λ^{-1}
median:	$I_{0.5}^{-1}(\alpha,\beta)$ no closed form	$\lambda^{-1}\ln 2$
mode:	$(\alpha-1)/(\alpha+\beta-2)$ for $\alpha>1, \beta>1$	0
variance:	$\alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1)$	λ^{-2}
skewness:	$\frac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$	2
ex. kurtosis:	see text	6
entropy:	see text	$1-\ln(\lambda)$
mgf:	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$	$\left(1-\frac{t}{\lambda}\right)^{-1}$
cf:	${}_1F_1(\alpha;\alpha+\beta;it)$	$\left(1-\frac{it}{\lambda}\right)^{-1}$

	DOUBLE - EXPONENTIAL DISTRIBUTION	LOGISTIC DISTRIBUTION
parameters:	μ location (real), $b > 0$ scale (real)	μ location (real) ; $s > 0$ scale (real)
support:	$x \in (-\infty ; +\infty)$	$x \in (-\infty ; +\infty)$
pdf:	$\frac{1}{2b} \exp\left(-\frac{\lvert x-\mu \rvert}{b}\right)$	$\frac{e^{-(x-\mu)/s}}{s\left(1+e^{-(x-\mu)/s}\right)^2}$
cdf:	see text	$\frac{1}{1+e^{-(x-\mu)/s}}$
mean:	μ	μ
median:	μ	μ
mode:	μ	μ
variance:	$2b^2$	$\frac{\pi^2}{3} s^2$
skewness:	0	0
ex. kurtosis:	3	$6/s$
entropy:	$\log(2eb)$	$\ln(s)+2$
mgf:	$\frac{\exp(\mu t)}{1-b^2t^2}$ for $\lvert t \rvert < 1/b$	$e^{\mu t} B(1-st, 1+st)$ for $\lvert st \rvert < 1$, Beta function
cf:	$\frac{\exp(\mu i t)}{1+b^2t^2}$	$e^{i\mu t} B(1-ist, 1+ist)$ for $\lvert ist \rvert < 1$.

	CAUCHY DISTRIBUTION	GAMMA DISTRIBUTION
parameters :	x_0 location (real) $\gamma > 0$ scale (real)	$k > 0$ shape $\theta > 0$ scale
support :	$x \in (-\infty, +\infty)$	$x \in [0, \infty)$
pdf :	$\dfrac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}$	$x^{k-1} \dfrac{\exp(-x/\theta)}{\Gamma(k)\,\theta^k}$
cdf :	$\dfrac{1}{\pi}\arctan\left(\dfrac{x - x_0}{\gamma}\right) + \dfrac{1}{2}$	$\dfrac{\gamma(k, x/\theta)}{\Gamma(k)}$
mean :	does not exist.	$k\theta$
median :	x_0	no simple closed form
mode :	x_0	$(k-1)\theta$ for $k \geq 1$
variance :	does not exist	$k\theta^2$
skewness :	does not exist	$\dfrac{2}{\sqrt{k}}$
ex. kurtosis :	does not exist	$\dfrac{6}{k}$
entropy :	$\log(\gamma) + \log(4\pi)$	$k + \ln\theta + \ln\Gamma(k) + (1-k)\,\psi(k)$
m.g.f. :	does not exist	$(1 - \theta t)^{-k}$ for $t < \frac{1}{\theta}$
c.f. :	$\exp\left(x_0 it - \gamma\,\lvert t\rvert\right)$	$(1 - \theta it)^{-k}$

	NORMAL DISTRIBUTION	UNIFORM DISTRIBUTION
notation :	$N(\mu, \sigma^2)$	$\mathcal{U}(a,b)$
parameters :	$\mu \in \mathbb{R}$ — mean (location) $\sigma^2 > 0$ — variance (squared scale)	$-\infty < a < b < \infty$
support :	$x \in \mathbb{R}$	$x \in [a,b]$
pdf :	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{OW} \end{cases}$
cdf :	$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right]$	$\begin{cases} 0 & \text{for } x \le a \\ \frac{x-a}{b-a} & \text{for } x \in [a,b] \\ 1 & \text{for } x \geq b \end{cases}$
mean :	μ	$\frac{1}{2}(a+b)$
median :	μ	$\frac{1}{2}(a+b)$
mode :	μ	any value in $[a,b]$
variance :	σ^2	$\frac{1}{12}(b-a)^2$
skewness :	0	0
ex. kurtosis :	0	$-\frac{6}{5}$
entropy :	$\frac{1}{2}\ln(2\pi e\sigma^2)$	$\ln(b-a)$
mgf :	$\exp\{\mu t + \frac{1}{2}\sigma^2 t^2\}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
cf :	$\exp\{i\mu t - \frac{1}{2}\sigma^2 t^2\}$	$\frac{e^{itb}-e^{ita}}{it(b-a)}$
Fisher information :	$\begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 1/(2\sigma^4) \end{pmatrix}$	

	PARETO DISTRIBUTION	LOG-NORMAL DISTRIBUTION
parameters:	$\alpha > 0$ shape (real)	$\sigma^2 > 0$ - shape (real), $\mu \in \mathbb{R}$ - log-scale
support:	$x \in [x_m ; +\infty)$	$x \in (0, +\infty)$
pdf:	$\frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ for $x > x_m$	$\frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$
cdf:	$1 - \left(\frac{x_m}{x}\right)^\alpha$	$\frac{1}{2} + \frac{1}{2}\operatorname{erf}\left[\frac{\ln x - \mu}{\sqrt{2\sigma^2}}\right]$
mean:	$\frac{\alpha x_m}{\alpha - 1}$ for $\alpha > 1$	$e^{\mu + \sigma^2/2}$
median:	$x_m \sqrt[\alpha]{2}$	e^{μ}
mode:	x_m	$e^{\mu - \sigma^2}$
variance:	$\frac{x_m^2 \alpha}{(\alpha-1)^2 (\alpha-2)}$ for $\alpha > 2$	$\left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$
skewness:	$\frac{2(1+\alpha)}{\alpha - 3}\sqrt{\frac{\alpha-2}{\alpha}}$ for $\alpha > 3$	$\left(e^{\sigma^2} + 2\right)\sqrt{e^{\sigma^2} - 1}$
ex. kurtosis:	$\frac{6(\alpha^3 + \alpha^2 - 6\alpha - 2)}{\alpha(\alpha-3)(\alpha-4)}$ for $\alpha > 4$	$e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$
entropy:	$\ln\left(\frac{x_m}{\alpha}\right) + \frac{1}{\alpha} + 1$	$\frac{1}{2} + \frac{1}{2}\ln(2\pi\sigma^2) + \mu$
mgf:	$\alpha(-x_m t)^\alpha \Gamma(-\alpha, -x_m t)$ for $t < 0$	(defined only on the negative half-axis, see text)
cf:	$\alpha(-i x_m t)^\alpha \Gamma(-\alpha, -i x_m t)$	representation $\sum_{n=0}^{\infty} \frac{(it)^n}{n!} \cdot e^{n\mu + n^2\sigma^2/2}$ is asymptotically divergent but sufficient for numerical purposes.

- IMPORTANT DISCRETE DISTRIBUTIONS discussed here:

1. Bernoulli's Distribution,
2. Binomial Distribution,
3. Negative Binomial Distribution,
4. Poisson Distribution,
5. Geometric Distribution,
6. Hyper-Geometric Distribution,
7. Power-Series Distribution,
8. Polya's Distribution,
9. Logarithmic Distribution.

Theoretical Distribution :— By the term 'theoretical distribution', we mean the distribution derived theoretically, i.e. without any empirical sense before hand we develop some special form of a random variable which may be discrete or continuous. Now, question arises, whether the distribution of this developed r.v. has any type of resemblance with any natural phenomena and the answer is 'YES'. For every theoretical distribution, there is a natural phenomena for which we develop the PMF or PDF of the r.v. and this procedure is called 'modelling' and some theoretical distn. are derived as by-product of some other theoretical distribution.

Example:— $f(x) = \frac{1}{\theta} e^{-x/\theta}, \; x>0$

We all know that testing of life-hour of a bulb is a destructive process. By the above ~~mode~~ PDF, we model the life hour distribution of the bulbs.

SOME DISCRETE DISTRIBUTIONS

BERNOULLI'S DISTRIBUTION :—

X is said to be a Bernoulli random variable with parameter p, $0 < p < 1$ if the p.m.f. of X is

$$f(x) = P[X = x] = p^x (1-p)^{1-x} \, I_x(0,1)$$

i.e.
$$f(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$$

Bernoulli Trials :— A set of trials is said to be a set of Bernoulli trials if, —

i) the result of any trials can be classified only under two categories, namely success and failure.

ii) the probability of success remains same in each trials.

iii) Trials are independent.

BINOMIAL DISTRIBUTION :—

A discrete random variable X with mass points $0, 1, 2, \ldots, n$ is said to follow Binomial distribution with parameters n and p if its p.m.f. is given by, —

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad , \; x = 0, 1, \ldots, n; \quad 0 < p < 1; \quad q = 1-p.$$

We write $X \sim Bin(n, p)$.

Result: If X denotes the number of success in a set of n Bernoullia trials with probability of success p per trial, then $X \sim Bin(n, p)$.

Proof:- Let,

A_i:- the event that the ith trial results in a success, $i = 1(1)n$.

The events $A_1, A_2, \ldots, A_n$ are independent and with probability $P(A_i) = p \;\; \forall \; i = 1(1)n$.

The mass-points of X are $0, 1, 2, \ldots, n$. If x be any such mass-points then,

$$P[X = x] = \binom{n}{x} P(A_1 A_2 \cdots A_x A^c_{x+1} A^c_{x+2} \cdots A^c_n)$$

$$= \binom{n}{x} P(A_1) P(A_2) \cdots P(A_x) P(A^c_{x+1}) P(A^c_{x+2}) \cdots P(A^c_n)$$

[$\because$ A_i's are independent]

$$= \binom{n}{x} p^x (1-p)^{n-x}$$

The p.m.f. of X is $f(x) = \binom{n}{x} p^x q^{n-x}$; $x = 0, 1, \ldots, n$; $0 < p < 1$; $p + q = 1$

i.e. $f(x) = P[\underbrace{SS\ldots S}_{x} \underbrace{FF\ldots F}_{n-x}] \binom{n}{x} = \binom{n}{x} p^x q^{n-x}$.

☆ Problem: 1. If $X \sim Bin(n,p)$. Find the distribution of $Y = n-X$.

Since mass points of X are $0,1,\ldots,n$.
Hence the mass points of Y are $0,1,\ldots,n$.

If y be any such mass point then,

$$P(Y=y) = P(n-X=y) = P(X=n-y) = \binom{n}{n-y} p^{n-y}(1-p)^y$$
$$= \binom{n}{y}(1-p)^y p^{n-y}$$

$\therefore n-X \sim Bin(n, 1-p)$

☆ 2. A box contains 2 coins with probability of heads 0·3 and 0·4, respectively. One of the coin is chosen at random and tossed n times. Find the probability distribution of the number of heads obtain.

Let X denotes the number of heads obtain.
Hence the mass points of X are $0,1,2,\ldots,n$.
A_i :- the event that the ith coin is chosen. If x be any mass-points of X then

$$P[X=x] = P[(X=x)\cap A_1] + P[(X=x)\cap A_2]$$
$$= P(A_1)P[(X=x)/A_1] + P[A_2].P[(X=x)|A_2]$$
$$= \frac{1}{2}\binom{n}{x}(\cdot 3)^x(\cdot 7)^{n-x} + \frac{1}{2}\binom{n}{x}(\cdot 4)^x(\cdot 6)^{n-x}$$

Factorial Moments :— Let $X \sim Bin(n,p)$, then the rth order factorial moment of X is given by, —

$$\mu_{[r]} = E[(X)_r] = E[X(X-1)\cdots(X-r+1)]$$
$$= \sum_{x=0}^{n}(x)_r f(x)$$
$$= \sum_{x=0}^{n} x(x-1)\cdots(x-r+1)\frac{n!}{x!(n-x)!}p^x q^{n-x}$$
$$= p^r n(n-1)\cdots(n-r+1)\sum_{x=r}^{n}\binom{n-r}{x-r}p^{x-r}q^{n-x}$$
$$= (n)_r p^r \sum_{y=0}^{n-r}\binom{n-r}{y}p^y q^{n-r-y}$$
$$= (n)_r p^r (p+q)^{n-r}$$
$$= (n)_r p^r$$

$$\therefore \mu_{[r]} = \begin{cases}(n)_r p^r & \text{if } r\le n \\ 0 & \text{ow}\end{cases}$$

Putting $r=1$, we get —

$$\mu_{[1]} = \mu_1' = E(X) = np$$

Putting $r=2$, we get —

$$\mu_{[2]} = E[X(X-1)] = n(n-1)p^2$$

$$\therefore Var(X) = E(X^2) - E^2(X) = E[X(X-1)] + E(X) - E^2(X)$$
$$= n(n-1)p^2 + np - n^2p^2$$
$$= np(1-p) = npq .$$

So, we can conclude that if $X \sim Bin(n,p)$ then $\boxed{\text{mean} > \text{Variance}}$ as $(p+q=1)$, $0<p<1$, i.e. $np > npq$.

☆3. If $X \sim Bin(n,p)$ then show that $V(X) \le \frac{n}{4}$.

$V(X) = npq$
$= np(1-p)$; differentiating w.r.t. p, we get —

$$\therefore \frac{dV(X)}{dp} = n - 2np = 0$$
$$\Rightarrow p = \frac{1}{2}$$

Now, $\frac{d^2V(X)}{dp^2} = -2n < 0$

[OR AM $\geq$ GM

$\Rightarrow \frac{p+(1-p)}{2} \geq \sqrt{p(1-p)}$

$\Rightarrow \frac{1}{2} \geq \sqrt{p(1-p)}$

$\Rightarrow \frac{n}{4} \geq npq$ as $n>0$]

$\therefore V(X)$ has a maxima at $p=\frac{1}{2}$.

$\therefore V(X)$ is maximum when $V(X) = n \cdot \frac{1}{2}\left(1-\frac{1}{2}\right) = \frac{n}{4}$.

$$\therefore V(X) \le \frac{n}{4}.$$

☆4. If $X \sim Bin(n,p)$, find $E\left(\frac{1}{X+1}\right)$ & $E\left[\frac{1}{(X+1)(X+2)}\right]$

$$E\left(\frac{1}{X+1}\right) = \sum_{x=0}^{n} \frac{1}{1+x}\binom{n}{x} p^x q^{n-x}$$
$$= \sum_{x=0}^{n} \frac{1}{(x+1)} \cdot \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$
$$= \frac{1}{p(n+1)} \sum_{x=0}^{n} \frac{(n+1)!}{(x+1)!\,(n-x)!} p^{x+1} q^{n-x}$$
$$= \frac{1}{p(n+1)} \sum_{x'=1}^{n} \binom{n+1}{x'} p^{x'} q^{n+1-x'} \qquad \left[\text{let, } (x+1) = x'\right]$$
$$= \frac{1}{p(n+1)} \left[\sum_{x'=0}^{n} \binom{n+1}{x'} p^{x'} q^{n+1-x'} - \binom{n+1}{0} p^0 q^{n+1}\right]$$
$$= \frac{1}{p(n+1)} \left[(p+q)^{n+1} - q^{n+1}\right]$$
$$= \frac{1}{p(n+1)} \left[1 - q^{n+1}\right]$$

$$E\left[\frac{1}{(X+1)(X+2)}\right] = \sum_{x=0}^{n} \frac{1}{(x+1)(x+2)} \binom{n}{x} p^x q^{n-x}$$

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

(Let, $x+2=y$),

$$= \frac{1}{p^2(n+2)(n+1)} \sum_{y=2}^{n+2} \binom{n+2}{y} p^y q^{n+2-y}$$

$$= \frac{1}{p^2(n+2)(n+1)} \left[\sum_{y=0}^{n+2} \binom{n+2}{y} p^y q^{n+2-y} - q^{n+2} - (n+2)pq^{n+1}\right]$$

$$= \frac{1}{p^2(n+2)(n+1)} \left[1 - q^{n+2} - (n+2)q^{n+1}p\right]$$

Raw moment: Let $X \sim Bin(n,p)$, then the rth order raw moment of X is given by,

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

$$= \sum_{x=0}^{n} x^{r-1} \cdot x \binom{n}{x} p^x q^{n-x}$$

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

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

[Let, $x' = x-1$]

$$= np\, E[1+x']^{r-1}, \text{ where } X' \sim Bin(n-1,p)$$

$\therefore \mu_r' = E(X^r)$ where $X \sim Bin(n,p)$

$= np\, E[1+x']^{r-1}$, where $X' \sim Bin(n-1,p)$

Putting $r=1$,

$\therefore \mu_1' = E(X) = np\, E[1+x']^0$

$= np$

Putting, $r=2$,

$\therefore \mu_2' = E(X^2) = np[E(1+x')] = np[1 + E(X')]$

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

$\therefore V(X) = np + np^2(n-1) - n^2p^2$

$= npq$

Central Moment : — (Recurrence Relation Regarding Moments)

Let $X \sim Bin(n, p)$, then the rth order central moment of X is given by, —

$$\Rightarrow \mu_r = E[X - E(X)]^r$$

$$\mu_r = E(X - np)^r$$

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

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

differentiating both side w.r.t. p, we get —

$$\frac{d\mu_r}{dp} = \sum_{x=0}^{n} \frac{d}{dp}\left[(x-np)^r \binom{n}{x} p^x (1-p)^{n-x}\right]$$

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

$$= -nr \sum_{x=0}^{n} \binom{n}{x} (x-np)^{r-1} p^x (1-p)^{n-x} + \sum_{x=0}^{n} \binom{n}{x} (x-np)^r p^{x-1} (1-p)^{n-x-1} [x(1-p) - (n-x)p]$$

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

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

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

$$\therefore \mu_{r+1} = pq\left[\frac{d\mu_r}{dp} + nr\mu_{r-1}\right];$$ [Renovsky formula]

(* It is called the Recursion Relation for central moments *)

Now, for $r = 1$,

$$\mu_2 = pq\left[\frac{d\mu_1}{dp} + n\mu_0\right] \quad [\text{as } \mu_0 = 1, \mu_1 = 0]$$

$$= npq$$

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

$$= pq\left[\frac{d}{dp}\{np(1-p)\} + 0\right]$$

$$= pq[n(1-p) - np]$$

$$= npq(1-2p) = npq(q-p)$$

Putting, $r = 3$, $\mu_4 = pq\left[\frac{d}{dp}\mu_3 + 3n\mu_2\right]$

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

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

Coefficient of skewness:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{n^2p^2q^2(1-2p)^2}{n^3p^3q^3} = \frac{(1-2p)^2}{npq} \gtreqless 0 \text{ according as } p \lesseqgtr q$$

$\therefore$ The distribution is symmetric if $(1-2p)=0 \Leftrightarrow p=\frac{1}{2}$ or, $p=q$

+vely skewed if $(1-2p)>0 \Leftrightarrow p<\frac{1}{2}$ or $p<q$

−vely skewed if $(1-2p)<0 \Leftrightarrow p>\frac{1}{2}$ or $p>q$

Intuitively, if $p=\frac{1}{2}$, $P(X=x) = \binom{n}{x}\left(\frac{1}{2}\right)^n = \binom{n}{n-x}\left(\frac{1}{2}\right)^n$

$= P(X=n-x) \ \forall\, x=0(1)n.$

$\therefore$ the resulting distribution is symmetric.

$p=q$ $\qquad$ $p<q$ $\qquad$ $p>q$

Coefficient of Kurtosis:

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{(1-6pq)}{npq}$$

Now, $\beta_2 \gtreqless 3$ according as $(1-6pq) \gtreqless 0$, i.e. $pq \lesseqgtr \frac{1}{6}$

$\therefore$ The distribution is mesokurtic when $(1-6pq)=0$, i.e. $pq=\frac{1}{6}$

leptokurtic when $(1-6pq)>0$, i.e. $pq<\frac{1}{6}$

platykurtic when $(1-6pq)<0$, i.e. $pq>\frac{1}{6}$

☆5. If $X \sim Bin(n,p)$, show that → i) $E\left(\frac{X}{n}-p\right)^2 = \frac{pq}{n}$,

ii) $Cov\left(\frac{X}{n}, \frac{n-X}{n}\right) = -\frac{pq}{n}$.

(i) $E\left(\frac{X}{n}-p\right)^2 = \frac{1}{n^2}E(X-np)^2 = \frac{1}{n^2}\cdot npq = \frac{pq}{n}$.

(ii) $Cov\left(\frac{X}{n}, \frac{n-X}{n}\right) = cov\left(\frac{X}{n}, 1-\frac{X}{n}\right) = -cov\left(\frac{X}{n},\frac{X}{n}\right) = -V\left(\frac{X}{n}\right) = -\frac{1}{n^2}\cdot npq$

$= -\frac{pq}{n}$

☆6. If X & Y independently follows Bin(n,p). Then find the distribution of (X−Y).

Mass points of X are $0,1,2,\ldots,n$ & same for Y.

Then the mass points of (X−Y) are $-n, -n+1, -n+2, \ldots, 0, 1, 2, \ldots, n-1, n$.

$$P(X-Y=Z) = \sum_{k=0}^{n} P[X=k, Y=k-z]$$

$$= \sum_{k=0}^{n} P(X=k)P(Y=k-z) \quad [\because X \,\&\, Y \text{ are independent}]$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \binom{n}{k-z} p^{k-z} q^{n-k+z}$$

$$= \sum_{k=0}^{n} \binom{n}{k}\binom{n}{k-z} p^{2k-z} q^{2n-2k+z} \quad ; z=-n(1)n.$$

Recurrence Relation Regarding Probabilities :

If $X \sim Bin(n, p)$

$$P(X=x) = \binom{n}{x} p^x q^{n-x} \;;\; x = 0(1)n, \; 0<p<1$$

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

so, $$\frac{P(X=x)}{P(X=x-1)} = \frac{\binom{n}{x}}{\binom{n}{x-1}} \cdot \frac{p}{q} = \frac{n-x+1}{x} \cdot \frac{p}{q}$$

$$\therefore P(X=x) = \frac{(n-x+1)p}{xq} \cdot P(X=x-1) \text{ ------ } (*)$$

4.29.

☆ 7. If $P[X=x] = P_x$ for a bi-nomially distributed random variable, say, with parameters (n, p), then show that

$$\frac{P_1}{P_0} \geq \frac{P_2}{P_1} \geq \frac{P_3}{P_2} \geq \cdots \geq \frac{P_n}{P_{n-1}}.$$

From the above $(*)$, we get, $$\frac{P(X=x)}{P(X=x-1)} = \frac{n-x+1}{x} \cdot \frac{p}{q}$$

$$= \left[\frac{n+1}{x} - 1\right]\frac{p}{q}$$

$\therefore \dfrac{P(X=x)}{P(X=x-1)}$ is a decreasing function of x.

$$\therefore \frac{P_1}{P_0} \geq \frac{P_2}{P_1} \geq \cdots \geq \frac{P_n}{P_{n-1}}.$$

☆ 8. Show Binomial variate as a sum of independent Bernoulli variates.

Suppose, $X_1, X_2, \ldots, X_n$ are independently distributed Bernoulli random variable with parameter p.

i.e. p.g.f. of X_i is

$$P(X=x_i) = p^x (1-p)^{1-x} I_x(0,1) \quad \forall\, x = 0, 1.$$

We know p.g.f. of $S_n = \sum_{i=1}^{n} X_i$,

$$P_{S_n}(t) = [P_{X_i}(t)]^n$$

Now, $$P_{X_i}(t) = E(t^{X_i}) = t^0(1-p) + tp$$
$$= 1 - p + tp.$$

$$\therefore P_{S_n}(t) = (1 - p + tp)^n$$
$$= \sum_{x=0}^{n} \binom{n}{x} (pt)^x (1-p)^{n-x}$$

$\therefore P[S_n = x] =$ coefficient of t^x in the expansion of $P_{S_n}(t)$.

$$= \binom{n}{x} p^x (1-p)^{n-x} \;;\; x = 0, 1, \ldots, n.$$

Mode of the Binomial Distribution :

Let $X \sim Bin(n, p)$, the p.m.f. of X is $f(x) = \binom{n}{x} p^x q^{n-x}$

$x = 0, 1, \ldots, n$

$0 < p < 1$, $p + q = 1$.

The mode of the distribution is the value(s) of x for which $f(x)$ is maximum. [mode = most probable value]

for $x = 1, 2, \ldots, n$.

Now,
$$\frac{f(x)}{f(x-1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x-1} p^{x-1} q^{n-x+1}} = \frac{(n-x+1)p}{xq}$$

Now,
$$\frac{f(x)}{f(x-1)} \gtreqless 1 \text{ according as } \frac{(n-x+1)p}{xq} \gtreqless 1$$

$$\Rightarrow np - xp + p \gtreqless x - xp \quad [\because q = 1-p]$$

$$\Rightarrow (n+1)p \gtreqless x$$

$\therefore f(x) \gtreqless f(x-1)$ according as $x \lesseqgtr (n+1)p$

Case-1: $\rightarrow$ $(n+1)p =$ integer $= K$ (say)

Now, for $x = K$,

$\Rightarrow f(K) = f(K-1) \quad \ldots\ldots (1)$

for $x < K$

$\Rightarrow f(x) > f(x-1)$ for $x = 1, \ldots, K-1$

$\Rightarrow f(0) < f(1) < \cdots < f(K-1) \quad \ldots\ldots (2)$

again for $x > K$

$\Rightarrow f(x) < f(x-1)$ for $x = K+1, K+2, \ldots, n$

$\Rightarrow f(K) > f(K+1) > \cdots > f(n) \quad \ldots\ldots (3)$

Combining (1), (2), (3); we get,

$$f(0) < f(1) < \cdots < f(K-1) = f(K) > f(K+1) > \cdots > f(n)$$

$\therefore$ $f(x)$ is maximum at $x = K$ and $x = K-1$

$\therefore$ the modes of the distribution are at $x = (n+1)p$ and $(n+1)p-1$

[e.g. [2.5] = 2]

(101) Case-2 : → $(n+1)p$ is not an integer, let $m = [(n+1)p]$
= largest integer obtained in $(n+1)p$.

Obviously, $m < (n+1)p < m+1$

Now, $x = m$

$\Rightarrow x < (n+1)p$

$\Rightarrow f(m) > f(m-1)$ ...... (4)

$x < m$

$\Rightarrow x < (n+1)p$

$\Rightarrow f(x) > f(x-1)$ for $x = 1, \ldots, m-1$

$\Rightarrow f_0 < f(1) < \ldots < f(m-1)$ .... (5)

$x > m$

$\Rightarrow x \geq m+1 > (n+1)p$

$\Rightarrow x > (n+1)p$

$\Rightarrow f(x) < f(x-1)$ for $x = m+1, \ldots, n$

$\Rightarrow f(m) > f(m+1) > \ldots > f(n)$ .... (6)

Combining (4), (5), (6); we get —

$f(0) < f(1) < f(2) < \ldots < f(m) > f(m+1) > \ldots > f(n)$

$\therefore f(x)$ is maximum at $x = m$,

$\therefore$ the mode of the distribution are at $x = (n+1)p$.

Example: → i) $X \sim Bin(12, \frac{1}{3})$; ii) $X \sim Bin(11, \frac{1}{3})$

$\therefore$ Mode $= \left[\frac{13}{3}\right] = 4$ $\therefore$ Mode $= \left[\frac{12}{3}\right] = 3, 4$.

☆9. If $X \sim Bin(n, p)$, find the restricted range of p for which there will be a single mode $x = 1$, where, $n = 21$.

$\therefore (n+1)p = (21+1)p = 22p$.

It is given that $[22p] = 1$, here $m = 1$ {by the property of box function}

$\therefore 1 \leq 22p < 2$

$\Rightarrow \frac{1}{22} \leq p < \frac{1}{11}$

☆ 10. If $X \sim Bin(2s, \frac{1}{2})$, s = integer, show that,

i) the most probable value of x is s.

ii) $P[X=s] = \frac{(2s-1)(2s-3)\ldots\ldots 3.1}{2s(2s-2)\ldots\ldots 4.2}$

iii) $\frac{1}{2\sqrt{s}} < P[X=s] < \frac{1}{\sqrt{2s+1}}$ [WBSU'11]

i) $(n+1)p = (2s+1)\frac{1}{2} = (s+\frac{1}{2})$

$\therefore [(n+1)p] = [s+\frac{1}{2}] = s$. Hence the proof.

ii) $P[X=s] = \binom{n}{s} p^{s}(1-p)^{n-s}$

$= \binom{2s}{s}\left(\frac{1}{2}\right)^{2s}$

$= \frac{1}{2^{2s}} \cdot \frac{2s!}{s!\,s!}$

$= \left(\frac{1}{2}\right)^{2s} \cdot \frac{2s(2s-1)(2s-2)\ldots\ldots 3.2.1}{\{s(s-1)\ldots 2.1\}\{s(s-1)\ldots 2.1\}}$

$= \frac{\cancel{2^s}(\cancel{s.\overline{s-1}\ldots 2.1})\{(2s-1)(2s-3)\ldots 3.1\}}{\cancel{2^s}.\cancel{s!}\{2s.(2s-2)(2s-4)\ldots 4.2\}}$

$= \frac{(2s-1)(2s-3)\ldots 3.1}{2s(2s-2)\ldots\ldots 4.2}$. Hence the proof.

iii) Consider the function, $g(y) = 1-\frac{1}{y}$, $y>1$

$\therefore g(y)$ is increasing function of y. Then —

$g(2k-1) < g(2k) < g(2k+1)$

$\therefore 1-\frac{1}{2k-1} < 1-\frac{1}{2k} < 1-\frac{1}{2k+1}$

$\Rightarrow \frac{2(k-1)}{2k-1} < \frac{2k-1}{2k} < \frac{2k}{2k+1}$

Putting $k = 1, 2, \ldots, s$ successively, we get

$\frac{1}{2} < \frac{2}{3}$

$\frac{3}{4} < \frac{4}{5}$

$\vdots$

$\frac{2s-1}{2s} < \frac{2s}{2s+1}$

Multiplying we get, $\frac{1.3\ldots(2s-1)}{2.4\ldots 2s} < \frac{2.4\ldots\ldots 2s}{3.5\ldots(2s+1)}$

$\Rightarrow P[X=s] < \frac{1}{P[X=s](2s+1)}$

$\Rightarrow P[X=s] < \frac{1}{\sqrt{2s+1}}$

Now, $\frac{2(k-1)}{2k-1} < \frac{2k-1}{2k}$

Putting $k = 2, 3, \ldots, s$; successively we get

$$\frac{2}{3} < \frac{3}{4}$$

$$\frac{4}{5} < \frac{5}{6}$$

$$\frac{6}{7} < \frac{7}{8}$$

$$\vdots$$

$$\frac{2(s-1)}{2s-1} < \frac{2s-1}{2s}$$

Multiplying we get,

$$\frac{2.4.6\ldots(2s-2)}{3.5.7\ldots(2s-1)} < \frac{1.3.5\ldots(2s-1)}{4.6\ldots 2s}$$

$$\Rightarrow \frac{1}{2 \cdot s\, P[X=s]} < 2P[X=s]$$

$$\Rightarrow \quad P[X=s] > \frac{1}{2\sqrt{s}}$$

$$\therefore \frac{1}{2\sqrt{s}} < P[X=s] < \frac{1}{\sqrt{2s+1}}$$

☆11. If $X \sim Bin(n,p)$. Show that $f(k) < \frac{1}{k-np}$ for $k > np$.

For, $k > np$, $\sum_{x=[np]+1}^{k} f(x) < 1$ [since its a probability]

since $f(x)$ is a decreasing function of x where $x > (n+1)p$.

$f(x) = f(x-1)$ when $x = (n+1)p$

$$(n+1)p \leq [np]+1$$

then $\quad 1 > \sum_{x=[np]+1}^{k} f(x) \geq \sum_{x=[np]+1}^{k} f(k)$

$$= f(k)\{k - [np]\}$$

$\therefore f(k) < \frac{1}{k-[np]} < \frac{1}{k-np}$ when $k > np$

☆ 12. If $X \sim Bin(n, p)$ then show that

i) $P[X \geq k] < f(k) \dfrac{kq}{k-np}$ if $k > np$

ii) $P[X \leq k] < f(k) \cdot \dfrac{(n-k)p}{np-k}$ if $k < np$

i) $f(x) = \binom{n}{x} p^x q^{n-x}$

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

$$\therefore \frac{f(k+1)}{f(k)} = \frac{n-k}{k+1} \cdot \frac{p}{q} < \frac{(n-k)p}{kq} = r \text{ (say)}$$

$$f(k+1) < r f(k)$$

Now, $$\frac{f(k+2)}{f(k+1)} = \frac{n-k-1}{k+2} \cdot \frac{p}{q} < r$$

$$\therefore f(k+2) < r f(k+1) < r^2 f(k)$$

similarly, $f(k+3) < r^3 f(k)$ and so on

Now, $$P[X \geq k] = f(k) + f(k+1) + \cdots + f(n)$$
$$= f(k) + f(k+1) + \cdots \infty$$
$$< f(k) + r f(k) + r^2 f(k) + \cdots \infty$$
$$= f(k) \cdot \frac{1}{1-r} \text{ if } |r| < 1$$
$$= f(k) \cdot \left[\frac{1}{1 - \frac{(n-k)}{kq} p}\right] \text{ if } |r| < 1 \text{ i.e. } \frac{(n-k)p}{kq} < 1$$
$$= f(k) \cdot \frac{kq}{k-np} \text{ if } k > np.$$

$\therefore P[X \geq k] < f(k) \cdot \dfrac{kq}{k-np}$ [Hence the proof]

ii)

$P[X \leq k]$

$= P[n - X \geq n - k]$

$= P[X' \geq k']$, where, $X' = n - X$, $X' \sim Bin(n, q)$
$k' = n - k$.

$g(x') = \binom{n}{x'} p^{x'} q^{n-x'}$, $x' = 0, 1, \ldots, n$.
$0 < q < 1$, $p = 1 - q$.

$\therefore g(k') = \binom{n}{k'} q^{k'} p^{n-k'}$

$= \binom{n}{n-k} q^{n-k} p^k = f(k)$

Now from the 1st part of the problem, we get

$$P[X' \geq K'] < g(K')\left[\frac{K'p}{K'-nq}\right] \text{ if } K' > nq$$

$$\Rightarrow P[X \leq K] < f(K)\left[\frac{(n-K)p}{np-K}\right] \text{ if } K < np.$$

Hence the proof.

☆13. If $X \sim Bin(n,p)$ then show that

i) $P[X \geq K] < \dfrac{Kq}{(K-np)^2}$ if $K > np$

ii) $P[X \leq K] < \dfrac{(n-K)p}{(np-K)^2}$ if $K < np$

Soln.

From the previous 2 problems, we get

$$P[X \geq K] < \frac{Kq}{(K-np)} f(K) \quad \cdots\cdots (1) \quad \text{where } K > np$$

$$P[X \leq K] < \frac{(n-K)p}{(np-K)} f(K) \quad \cdots\cdots (2) \quad \text{where } K < np$$

and

$$f(K) < \frac{1}{K-np} \quad \text{where } K > np \cdots\cdots (3)$$

$$< \frac{1}{np-K} \quad \text{where } K < np \cdots\cdots (4)$$

From (1) & (3), we get,

$$P[X \geq K] < \frac{Kq}{K-np} f(K) < \frac{Kq}{(K-np)^2} \quad \text{where } K > np.$$

& from (2) & (4), we get,

$$P[X \leq K] < \frac{(n-K)p}{np-K} f(K) < \frac{(n-K)p}{(np-K)^2} \quad \text{where } np > K$$

Hence the proof is complete.

OR, $P(X \geq K) = P[X - np \geq K - np] \leq P[\,|X-np| \geq K-np]$ if $K > np$

$$= P[\,|X - E(X)| \geq K-np]$$

$$\leq \frac{Var(X)}{(K-np)^2} = \frac{npq}{(K-np)^2}$$

$$< \frac{Kq}{(K-np)^2}$$

$\left[\because P[\,|X-\mu| \geq t] \leq \frac{\sigma^2}{t^2}\right.$, Chebyshev's inequality$\left.\right]$

$\cancel{f(K) < \frac{1}{K-np}, K > np}$

$\cancel{\text{so, } P(X \geq K) < \frac{Kq}{(K-np)^2}}$

$\cancel{< \frac{1}{(K-np)}\frac{Kq}{(K-np)}}$

Mean Deviation about Mean : [C.U.]

$$MD_\mu(X) = E|X-\mu|$$
$$= \sum_{x=0}^{n} |x-\mu| f(x)$$
$$= \sum_{x=0}^{[\mu]} (\mu - x) f(x) + \sum_{x=[\mu]+1}^{n} (x-\mu) f(x)$$
$$= 2\sum_{x=[\mu]+1}^{n} (x-\mu) f(x) \qquad [\text{ since } E(X-np)=0]$$
$$[m_0 = [\mu]+1,\ \mu = np]$$
$$= 2\sum_{x=m_0}^{n} (x-np) p^x q^{n-x} \binom{n}{x}$$
$$= 2\left[\sum_{x=m_0}^{n} x \cdot \frac{n}{x}\binom{n-1}{x-1} p^x q^{n-x} - np\sum_{x=m_0}^{n}\left\{\binom{n-1}{x}+\binom{n-1}{x-1}\right\} p^x q^{n-x}\right]$$
$$= 2np\left[\sum_{x=m_0}^{n}\binom{n-1}{x-1} p^{x-1} q^{n-x} - \sum_{x=m_0}^{n-1}\binom{n-1}{x} p^x q^{n-x} - \sum_{x=m_0}^{n}\binom{n-1}{x-1} p^x q^{n-x}\right]$$
$$= 2np\left[\sum_{x=m_0}^{n}\binom{n-1}{x-1} p^{x-1} q^{n-x}(1-p) - \sum_{x=m_0}^{n-1}\binom{n-1}{x} p^x q^{n-x}\right]$$
$$= 2np\left[\sum_{x'=m_0-1}^{n-1}\binom{n-1}{x'} p^{x'} q^{n-x'} - \sum_{x=m_0}^{n-1}\binom{n-1}{x} p^x q^{n-x}\right] \quad [\because x' = x-1]$$
$$= 2np\binom{n-1}{m_0-1} p^{m_0-1} q^{n-m_0+1}$$
$$= 2m_0 q \binom{n}{m_0} p^{m_0} q^{n-m_0} = 2m_0 q\, P[X=m_0]$$
$$\simeq 2m_0 q \frac{n!}{m_0!\,(n-m_0)!} p^{m_0} q^{n-m_0}$$

Stirling approximation formula :– $n! \simeq \sqrt{2\pi}\, e^{-n} n^{n+\frac{1}{2}}$,

$$\therefore MD_\mu(X) = \frac{2m_0 q \cancel{\sqrt{2\pi}} \cdot \cancel{e^{-n}} \cdot n^{n+\frac{1}{2}} \cdot p^{m_0} q^{n-m_0}}{\cancel{\sqrt{2\pi}}\, \cancel{e^{-m_0}}\, m_0^{m_0+\frac{1}{2}} \sqrt{2\pi}\, \cancel{e^{-n+m_0}}\, (n-m_0)^{n-m_0+\frac{1}{2}}}$$
$$= \sqrt{\frac{2npq}{\pi}} \left(\frac{np}{m_0}\right)^{m_0-\frac{1}{2}} \left(\frac{nq}{n-m_0}\right)^{n-m_0+\frac{1}{2}}$$

for large n, $\frac{m_0}{n} \to p \quad \therefore \frac{n-m_0}{nq} \to 1$ & $\frac{np}{m_0} \to 1$.

$$\therefore E|X-\mu| = \sqrt{\frac{2npq}{\pi}} = \sqrt{\frac{2}{\pi}}\sqrt{V(X)}$$

$$\therefore \boxed{MD_\mu(X) \simeq \sqrt{\frac{2npq}{\pi}}} \qquad \therefore \boxed{\frac{MD_\mu(X)}{\sqrt{V(X)}} = \sqrt{\frac{2}{\pi}}} \quad [\text{Geari's Ratio}]$$

Distribution function in terms of Incomplete Beta function :

Incomplete Beta function,

$$I_x(m,n) = \frac{\int_0^x u^{m-1}(1-u)^{n-1}\,du}{\int_0^1 u^{m-1}(1-u)^{n-1}\,du} = \frac{\int_0^x u^{m-1}(1-u)^{n-1}}{B(m,n)} \; ; \; m>0,\, n>0$$

P.d.f. of Beta distn. of 1st kind is $f(x) = \begin{cases} \dfrac{x^{m-1}(1-x)^{n-1}}{B(m,n)} & \text{if } 0<x<1 \\ 0 & \text{o/w} \end{cases}$

Let, $X \sim Bin(n,p)$ then show that,

$$P[X \le k] = I_q(n-k, k+1).$$

Proof :-

$$P[X \le k] = \sum_{j=0}^{k} \binom{n}{j} p^j q^{n-j}$$

$$= \frac{1}{\frac{(n-k-1)!\,k!}{n!}} \sum_{j=0}^{k} \frac{n!}{j!\,(n-j)!} \cdot \frac{k!\,(n-k-1)!}{n!} p^j q^{n-j}$$

$$= \frac{1}{\frac{\Gamma(n-k)\Gamma(k+1)}{\Gamma(n+1)}} \sum_{j=0}^{k} \frac{k!}{j!\,(k-j)!} \frac{(k-j)!\,(n-k-1)!}{(n-j)!} p^j q^{n-j}$$

$$= \frac{1}{B(n-k,k+1)} \sum_{j=0}^{k} \binom{k}{j} \frac{\Gamma(k-j+1)\Gamma(n-k)}{\Gamma(n-j+1)} p^j q^{n-j}$$

$$= \frac{1}{B(n-k,k+1)} \sum_{j=0}^{k} \binom{k}{j} B(n-k, k-j+1) p^j q^{n-j}$$

$$= \frac{1}{B(n-k,k+1)} \sum_{j=0}^{k} \binom{k}{j} \left[\int_0^1 u^{n-k-1} (1-u)^{k-j} du \right] p^j q^{n-j}$$

$$= \frac{1}{B(n-k,k+1)} \int_0^1 \left[\sum_{j=0}^{k} \binom{k}{j} \cdot p^j \{q(1-u)\}^{k-j} \right] q^{n-k} \cdot u^{n-k-1} du$$

$$= \frac{1}{B(n-k,k+1)} \int_0^1 [p + q(1-u)]^k \cdot (qu)^{n-k-1} \cdot q\,du$$

$[p+q=1]$ [Let $qu = t$, $q\,du = dt$]

$$= \frac{1}{B(n-k,k+1)} \int_0^q (1-t)^{(k+1)-1} t^{(n-k)-1} dt$$

$$= I_q(n-k, k+1)$$

Hence, the proof.

Now, $I_x(m,n) = \left\{ \int_0^x u^{m-1}(1-u)^{n-1} du \right\} \Big/ B(m,n)$

$$= \frac{\int_0^1 u^{m-1}(1-u)^{n-1} du - \int_x^1 u^{m-1}(1-u)^{n-1} du}{B(m,n)}$$

Let, $1-u = t$, $\therefore -du = dt$

$$= 1 + \frac{\int_{1-x}^{0} (1-t)^{m-1} t^{n-1} dt}{B(m,n)}$$

$$= 1 - \frac{\int_0^{1-x} t^{n-1} (1-t)^{m-1} dt}{B(m,n)} = 1 - I_{1-x}(n,m)$$

From the above relationship, it could be said that,

$$I_q(n-k, k+1) = 1 - I_p(k+1, n-k).$$

Alternitive way:- If $X \sim Bin(n,p)$ then for any integer k, $0 \le k \le n$, $P[X \le k] = I_q(n-k, k+1)$, where,

$$I_s(a,b) = \frac{\int_0^s t^{a-1}(1-t)^{b-1}\,dt}{\int_0^1 t^{a-1}(1-t)^{b-1}\,dt} \quad \text{[incomplete Beta function]}, \; 0<s<1$$

C.U.

Ans:- $$P[X \le k] = \sum_{x=0}^{k} f(x)$$

$$= \sum_{x=0}^{k} \binom{n}{x} (1-q)^x q^{n-x} = G(q) \cdots\cdots \text{(say)}$$

Differentiating both sides w.r.t. q, we get

$$G'(q) = \sum_{x=0}^{k} \binom{n}{x} \left[x(1-q)^{x-1}(-1)q^{n-x} + (1-q)^x (n-x) q^{n-x-1} \right]$$

$$= \sum_{x=0}^{k} \left[\frac{n!}{(n-x-1)!\,x!}(1-q)^x q^{n-x-1} - \frac{n!}{(x-1)!\,(n-x)!}(1-q)^{x-1} q^{n-x} \right]$$

$$= \sum_{x=0}^{k} n \left[\binom{n-1}{x}(1-q)^x q^{n-x-1} - \binom{n-1}{x-1}(1-q)^{x-1} q^{n-x} \right]$$

$$= n\sum_{x=0}^{k} \binom{n-1}{x} p^x q^{n-x-1} - n\sum_{x=1}^{k} \binom{n-1}{x-1} p^{x-1} q^{n-x}$$

$$= n\sum_{x=0}^{k} \binom{n-1}{x} p^x q^{n-x-1} - n\sum_{x'=0}^{k-1} \binom{n-1}{x'} p^{x'} q^{n-x'-1}$$

[Let, $(x-1) = x'$]

$$= n\binom{n-1}{k} p^k q^{n-k-1}$$

$$= \frac{n!}{k!\,(n-k-1)!}(1-q)^k q^{n-k-1}$$

$$= \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k)}(1-q)^k q^{n-k-1}$$

$$= \frac{1}{B(n-k, k+1)}(1-q)^k q^{n-k-1}$$

$$\therefore G'(t) = \frac{(1-t)^k t^{n-k-1}}{B(n-k, k+1)}$$

$$\therefore G(q) = P[X \le k] = \int_0^q G'(t)\,dt = \frac{\int_0^q (1-t)^k t^{n-k-1}\,dt}{B(n-k, k+1)}$$

$$= I_q(n-k, k+1)$$

Distribution function of X is an increasing function of q and decreasing function of p.

☆ 14. If $X \sim Bin(n,p)$ Prove that

$$P[X \geq k] < \binom{n}{k} p^k \quad , k = 0, 1, \ldots, n.$$

ANS :-

$$P[X \leq k] = I_q(n-k, k+1)$$

$$P[X > k] = 1 - I_q(n-k, k+1)$$

$$= I_p(k+1, n-k)$$

$$\therefore P[X \geq k] = P[X > k-1]$$

$$= I_p(k, n-k+1)$$

$$= \frac{\int_0^p t^{k-1}(1-t)^{n-k}\,dt}{B(k, n-k+1)} < \frac{\int_0^p t^{k-1}\,dt}{B(k, n-k+1)}, \text{ since } 1 > (1-t)^{n-k}$$

$$= \left[\frac{t^k}{k}\right]_0^p \times \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)}$$

$$= \frac{p^k}{k} \times \frac{n!}{(k-1)!\,(n-k)!}$$

$$= p^k . \binom{n}{k} .$$

☆ 15. If $X \sim Bin(n,p)$. Find the mean.

ANS:-

$$\mu = E(X) = \sum_{x=0}^{n} x f(x)$$

$$= \sum_{x=0}^{n} x . \frac{n!}{x!\,(n-x)!} p^x q^{n-x} \text{ ; where } p+q=1.$$

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

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

$$= np \sum_{x'=0}^{n} \frac{(n-1)!}{x'!\,(n-x'-1)!} p^{x'} q^{n-x'-1} \quad [\text{where, } x' = x-1]$$

$$= np\,[p+q]^{n-1}$$

$$= np .$$

Scanned by CamScanner

Recurrence Relation for Cumulants :

If $X \sim Bin(n, p)$, then, moment generating function is $= (q+pe^t)^n$.

$$K(t) = \ln M(t)$$
$$= \ln (q+pe^t)^n$$
$$= n \ln (q+pe^t)$$
$$= n \ln \{q(1+\frac{p}{q}e^t)\}$$
$$= n \ln q + n \ln (1+\frac{p}{q}e^t)$$
$$= n \ln q + n \ln (1+e^{t+z}) \quad \text{Let, } z = \ln \frac{p}{q}$$

and $F(t+z) = n \ln (1+e^{t+z})$

$$\therefore K(t) = n \ln q + F(z) + tF'(z) + \frac{t^2}{2!}F''(z) + \cdots \infty$$
$$= \sum_{r=1}^{\infty} K_r \frac{t^r}{r!} \quad [\because n \ln q + F(z) = 0]$$

$$\therefore K_{r+1} = \frac{d^{r+1}}{dz^{r+1}} F(z)$$
$$= \frac{d}{dz}\left[\frac{d^r}{dz^r} F(z)\right]$$
$$= \frac{d}{dz} K_r = \frac{d}{dp} K_r \cdot \frac{dp}{dz} = pq \cdot \frac{d}{dp} K_r .$$

[Now, $z = \ln\left(\frac{p}{q}\right) = \ln p - \ln q$

$$\frac{dz}{dp} = \frac{1}{p} + \frac{1}{q} = \frac{1}{pq} \Rightarrow \frac{dp}{dz} = pq \,]$$

$$\therefore K_r = \frac{d^r}{dz^r} F(z) = \frac{d^{r-1}}{dz^{r-1}}\left(\frac{d}{dz}F(z)\right)$$
$$= \frac{d^{r-1}}{dz^{r-1}}\left[\frac{d}{dp}F(z) \cdot \frac{dp}{dz}\right]$$

Note that, $F(z) + n \ln q = 0$

$$\Rightarrow F(z) = -n \ln q$$
$$\Rightarrow \frac{d}{dp}(F(z)) = \frac{n}{q}$$

$$\therefore K_r = \frac{d^{r-1}}{dz^{r-1}}\left[\frac{n}{q} \cdot pq\right] = \frac{d^{r-1}}{dz^{r-1}}(np)$$

Putting $r = 1$, $\Rightarrow K_1 = np = \mu_1'$

$r = 2$, $\Rightarrow K_2 = pq \frac{d}{dp} K_1$

$$= pq \frac{d}{dp}(np) = npq = \mu_2$$

$r=3, \Rightarrow K_3 = pq \cdot \frac{d}{dp}(K_2)$

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

$= npq(q-p) = \mu_3$

$r=4 \Rightarrow K_4 = pq \frac{d}{dp}(K_3)$

$= npq \frac{d}{dp}[pq(q-p)]$

$= npq(1-6pq)$

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

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

☆ 16. If $X \sim Bin(n,p)$ then show that $\lim\limits_{n\to\infty} M_Z(t) = e^{t^2/2}$, where $Z = \frac{x-np}{\sqrt{npq}}$

<u>Ans:-</u>

$$M_Z(t) = E\left(e^{tZ}\right)$$

$$= E\left[e^{t\left(\frac{X-np}{\sqrt{npq}}\right)}\right]$$

$$= e^{-t\sqrt{\frac{np}{q}}}\, E\left[e^{\frac{tX}{\sqrt{npq}}}\right]$$

$$= e^{-t\sqrt{\frac{np}{q}}}\, M_X\left(\frac{t}{\sqrt{npq}}\right)$$

We know, $M_X(t) = (q+pe^t)^n$.

$$\therefore M_Z(t) = e^{-t\sqrt{\frac{np}{q}}}\left(q+pe^{\frac{t}{\sqrt{npq}}}\right)^n$$

$$= \left\{e^{-t\sqrt{\frac{p}{nq}}}\left(q+pe^{\frac{t}{\sqrt{npq}}}\right)\right\}^n$$

$$= \left\{qe^{-t\sqrt{\frac{p}{nq}}} + pe^{\frac{t}{\sqrt{npq}} - t\sqrt{\frac{p}{nq}}}\right\}^n$$

$$= \left\{qe^{-t\sqrt{\frac{p}{nq}}} + pe^{\frac{t}{\sqrt{nq}}\left(\frac{1}{\sqrt{p}} - \sqrt{p}\right)}\right\}^n$$

$$= \left\{qe^{-t\sqrt{\frac{p}{nq}}} + pe^{t\sqrt{\frac{q}{np}}}\right\}^n$$

$$= \left\{q\left(1 - t\sqrt{\frac{p}{nq}} + \frac{t^2 p}{2nq} + O(n^{-3/2})\right) + p\left(1 + t\sqrt{\frac{q}{np}} + \frac{t^2 q}{2np} + O(n^{-3/2})\right)\right\}^n$$

$$= \left[1 + \frac{t^2}{2n} + O(n^{-3/2})\right]^n$$

$$\therefore \lim_{n\to\infty} M_Z(t) = \lim_{n\to\infty}\left(1+\frac{t^2}{2n}\right)^n = e^{t^2/2}$$

$f(n) = O(n)$ if $\frac{f(n)}{n} \to$ a non-zero finite quantity for a large n as $n\to\infty$. $\frac{f(n)}{n} \to 0$ as $n \to \infty$.

e.g. $f(n) = n + \frac{2}{n}$ $\therefore \frac{f(n)}{n} = 1 + \frac{2}{n^2} \to 1$

$\therefore f(n) = O(n)$; again $\frac{f(n)}{n^2} = \frac{1}{n} + \frac{2}{n^3} \to 0$, $\therefore f(n) = O(n^2)$

☆ **Problem No. 17.** A drunk performs a 'random walk' over positions $0, \pm 1, \pm 2, \ldots$, as follows:

18)

He starts at 0. He takes successive one-unit steps, going to the right with probability 'p' and to the left with probability $(1-p)$. His steps are independent. Let X denotes his position after n steps. Find the distribution of $\left(\frac{X+n}{2}\right)$ and then find $E(X)$.

Soln. :→ $X_i = \begin{cases} -1, & \text{if he steps to the left with probability } (1-p). \\ 1, & \text{if he steps to the right with probability } p. \end{cases}$

for ith step, $i = 1(1)n$.

Then, $\frac{X_i+1}{2} = \begin{cases} 0, & \text{if he steps to the left with prob. } (1-p) \\ 1, & \text{if " " " right " " } p. \end{cases}$

is a Bernoulli random variable, $i = 1(1)n$.

Here, X_i's are i.i.d. Bernoulli r.v.'s, $i = 1(1)n$, then,

$$X = \sum_{i=1}^{n} X_i$$

Note that → $\sum_{i=1}^{n}\left(\frac{X_i+1}{2}\right) \sim Bin(n, p)$.

$$\Rightarrow \frac{\sum_{i=1}^{n} X_i + n}{2} \sim Bin(n,p) \Rightarrow \frac{X+n}{2} \sim Bin(n,p)$$

Here, $E\left(\frac{X+n}{2}\right) = np$

$$\Rightarrow E(X) + n = 2np$$

$$\Rightarrow E(X) = (2p-1)n.$$

☆ **Problem 18.** Let mutually independent r.v.'s X_1, X_2, X_3 have the same p.d.f. $f(x) = \begin{cases} 3x^2, & 0 < x < 1 \\ 0, & \text{o/w} \end{cases}$

Find the prob. that exactly 2 of these 3 variables exceeds 1/2.

Soln. :→ Consider $X_i > \frac{1}{2}$ as success, $i = 1(1)3$. X_i's are independent Bernoulli trials with $p = \sum_{i=1}^{3} [X_i > \frac{1}{2}] = \int_{1/2}^{1} 3x^2 dx = \frac{7}{8}$.

Define, Y = the no. of X_i's which are $> \frac{1}{2}$
= the no. of successes in 3 Bernoulli trials.
$\sim Bin(n = 3, p = \frac{7}{8})$

Required prob. $= P[Y = 2]$

$$= \binom{3}{2} p^2 q^{3-2} = \binom{3}{2}\left(\frac{7}{8}\right)^2\left(\frac{1}{8}\right)$$

$$= 3 \cdot \frac{7^2}{8^3} = \frac{147}{512}$$

NEGATIVE BINOMIAL DISTRIBUTION :
Or
INVERSE BINOMIAL DISTRIBUTION :

Negative Binomial distribution will be appropriate if the counts of an event which occurs in clusters such as,—

1) The distribution of mosquito bites.

2) The distribution of the number of eggs laid by an insect on leaves of a tree.

Definition :→ X is said to have a negative binomial distribution (or, inverse binomial distn. or waiting time binomial distribution) if the p.m.f. of X is of the form

$$P[X=x] = \binom{-r}{x} p^r (-q)^x \, I_x\{0,1,\ldots\}$$

when $0<p<1$ and $q=1-p$, $r \in \mathbb{N}$

We denote, $X \sim NB(r,p)$

$$\binom{-r}{x} = \frac{(-r)_x}{x!} = \frac{-r(-r-1)\cdots(-r-x+1)}{x!}$$

$$= \frac{(-1)^x (r+x-1)\cdots(r+1)r}{x!}$$

$$= \frac{(r+x-1)_x}{x!}(-1)^x$$

$$= \binom{r+x-1}{x}(-1)^x$$

i.e. $P[X=x] = \binom{x+r-1}{x} p^r q^x \, I_x\{0,1,\ldots\}$

[We know, $p^r (1-q)^{-r} = 1$

$$\Rightarrow \sum_{x=0}^{\infty} p^r \frac{(-r)_x}{x!}(-q)^x = 1$$

i.e. $$\sum_{x=0}^{\infty} \binom{r+x-1}{x} p^r q^x = 1.$$

So, it's a p.m.f.]

* The mases of negative binomial distribution over the different points are obtained as the different terms in the expansion of

$$p^r (1-q)^{-r} = \left(\frac{1}{p} - \frac{q}{p}\right)^{-r}$$

here as the index is a negative integer, so the distribution is named as <u>negative binomial distribution</u>.

* The distribution is also called a Waiting time binomial distribution because the mass points of x can be obtained from a sequence of Bernoulli trials with a stopping rule.

:→ Consider a sequence of Bernoulli trials with success probability p. Let the trials be repeated till the occurance of 'r' success.

Define X = No. of failures preceding the rth success.

$P[X=x] = P[$ first $[x+(r-1)]$ trials result in $(r-1)$ success and $(r+x)$th trial results in a success$]$

$$= \binom{x+r-1}{r-1} p^{r-1} q^{x} \cdot p$$

$$= \binom{x+r-1}{x} p^{r} q^{x}\, I_x\{0,1,\ldots\infty\}$$

since the event $\{X=x\}$ may occur in $\frac{(x+r-1)!}{x!\,(r-1)!}$ mutually exclusive and exhaustive and equally probable ways.

Probability Model: ⤳ Consider an indefinite series of Bernoulli trials. Let, $P\ (0<P<1)$ denotes the probability of a success in a trial. Let the trials be repeated until we are getting 'r' success where r is a pre-fixed positive integer. Let,

Y = No. of trials required to get r successes.

Mass points of Y are $r, r+1, r+2, \ldots\ldots \infty$. If y be any such mass point then,

$P[Y=y] = P[$ exactly y trials required to get r success$]$

$= P[$ In first $(y-1)$ trials we get $(r-1)$ successes and the yth trial results a success$]$

$$= \binom{y-1}{r-1} p^{r-1} q^{y-r} \cdot p$$

$$= \binom{y-1}{r-1} p^{r} q^{y-r}\, I_y\{r, r+1, \ldots\}$$

$\boxed{\text{OR}}$

* $P[$a particular case$]$

$$= P[\underbrace{SS\ldots\ldots S}_{(r-1)}\ \underbrace{FF\ldots\ldots F}_{(y-r)}S]$$

The particular case occur $\binom{y-1}{r-1}$ times.

$$P[Y=y] = \binom{y-1}{r-1} p^{r-1} \cdot q^{y-r} \cdot p$$

$$= \binom{y-1}{r-1} p^{r} q^{y-r}$$

Alternatively,

Now, if X denotes the number of failures preceeding the rth success.

$$\therefore Y = r + X \Rightarrow X = Y - r$$

mass points of X are $0, 1, 2, \ldots \infty$.

If x be any such mass points then

$$P[X=x] = P[Y-r=x] = P[Y=r+x] = \binom{r+x-1}{x} p^r q^x$$

$$= \binom{-r}{x} (-1)^x p^r q^x$$

$$= \binom{-r}{x} p^r (-q)^x$$

Factorial Moments : — Let $X \sim NB(r, p)$.

The Kth order factorial moment is

$$\mu_{[K]} = E[(X)_K] = \sum_{x=0}^{\infty} (x)_K f(x) = \sum_{x=0}^{\infty} (x)_K \binom{x+r-1}{x} p^r q^x$$

$$= \sum_{x=K}^{\infty} \frac{(x+r-1)!}{(x-K)!(r-1)!} p^r q^x$$

$$= q^K \frac{(K+r-1)!}{(r-1)!} \sum_{x=K}^{\infty} \frac{(r+x-1)!\, p^r q^{x-K}}{(x-K)!(r+K-1)!}$$

$$= q^K \binom{K+r-1}{K} K! \sum_{y=0}^{\infty} \frac{(r+y+K-1)!}{y!(r+K-1)!} p^r q^y$$

$$= q^K (K+r-1)_K \, p^r (1-q)^{-(r+K)} \quad [\because x-K=y \text{ (let)}]$$

$$= (K+r-1)_K \left[\frac{q}{p}\right]^K$$

$$\mu_{[1]} = (r)_1 \left(\frac{q}{p}\right)$$

$$= r \cdot \frac{q}{p}$$

$$\mu_{[2]} = (r+1)_2 \left(\frac{q}{p}\right)^2 = (r+1)r \left(\frac{q}{p}\right)^2$$

$$\mu_2 = \mu_{[2]} - \mu_{[1]}^2 + \mu_{[1]}$$

$$= r(r+1)\frac{q^2}{p^2} - r^2\frac{q^2}{p^2} + r\frac{q}{p}$$

$$= r \cdot \frac{q}{p^2}$$

$\boxed{\therefore \text{mean} < \text{variance.}}$

Scanned by CamScanner

Distribution function in terms of Incomplete Beta function :-

Result:- If $X \sim NB(r, p)$ then for any non-negative integer k,

$$P[X \le k] = I_p(r, k+1) \quad \text{, where}$$

$$I_s(a,b) = \frac{1}{B(a,b)} \int_0^s t^{a-1}(1-t)^{b-1}\,dt$$

Proof:-

$$P[X \le k] = \sum_{x=0}^{k} \binom{r+x-1}{x} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} B(r,k+1) \binom{r+x-1}{x} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \frac{\Gamma(r)\Gamma(k+1)}{\Gamma(r+k+1)} \binom{r+x-1}{x} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \frac{\cancel{(r-1)!}\,(k)!}{(r+k)!} \cdot \frac{(r+x-1)!}{x!\,\cancel{(r-1)!}} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \binom{k}{x} \frac{(k-x)!\,(r+x-1)!}{(r+k)!} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \binom{k}{x} \frac{\Gamma(k-x+1)\Gamma(r+x)}{\Gamma(r+k+1)} p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \binom{k}{x} B(r+x, k-x+1) p^r q^x$$

$$= \frac{1}{B(r,k+1)} \sum_{x=0}^{k} \binom{k}{x} \int_0^1 u^{r+x-1}(1-u)^{k-x}\,du\, p^r q^x$$

$$= \frac{1}{B(r,k+1)} \int_0^1 p^r u^{r-1} \sum_{x=0}^{k} \binom{k}{x} (qu)^x (1-u)^{k-x}\,du$$

$$= \frac{1}{B(r,k+1)} \int_0^1 p^r u^{r-1} (qu+1-u)^k\,du$$

$$= \frac{1}{B(r,k+1)} \int_0^1 (pu)^{r-1} (1-pu)^k p\,du$$

Let, $pu = t$
$\Rightarrow p\,du = dt$
[differentiating]

$$= \frac{1}{B(r,k+1)} \int_0^p t^{r-1}(1-t)^k\,dt$$

$$= I_p(r, k+1).$$

Alternative way :-

$$P[X \le k] = \sum_{x=0}^{k} \binom{r+x-1}{x} p^{r} (1-p)^{x} = G(p) \; \ldots \text{ say}$$

Differentiating $G(p)$ w.r.t. p, we get —

$$G'(p) = \sum_{x=0}^{k} \binom{r+x-1}{x} \left\{ r p^{r-1} (1-p)^{x} - p^{r} x (1-p)^{x-1} \right\}$$

$$= \sum_{x=0}^{k} \left[\binom{r+x-1}{x} p^{r-1} r (1-p)^{x} - \frac{(r+x-1)!}{r!\,(x-1)!} r p^{r} (1-p)^{x-1} \right]$$

$$= \sum_{x=0}^{k} r \left[\binom{r+x-1}{x} p^{r-1} (1-p)^{x} - \binom{r+x-1}{x-1} p^{r-1} (1-p)^{x-1} (1-\overline{1-p}) \right]$$

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

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

$$= \sum_{x=0}^{k} r \left[\binom{r+x-1}{x} + \binom{r+x-1}{x-1} \right] p^{r-1} (1-p)^{x} - \sum_{x=0}^{k} r \binom{r+x-1}{x-1} p^{r-1} (1-p)^{x-1}$$

$$= \sum_{x=0}^{k} r \left[\binom{r+x}{x} p^{r-1} (1-p)^{x} \right] - \sum_{x=1}^{k} r \left[\binom{r+x-1}{x-1} p^{r-1} (1-p)^{x-1} \right]$$

[Let $x-1=x'$ in the second summation]

$$= \sum_{x=0}^{k} r \binom{r+x}{x} p^{r-1} (1-p)^{x} - \sum_{x'=0}^{k-1} r \binom{r+x'}{x'} p^{r-1} (1-p)^{x'}$$

$$= \binom{r+k}{k} r p^{r-1} (1-p)^{k}$$

$$= \frac{(r+k)!}{k!\,(r-1)!} p^{r-1} (1-p)^{k}$$

$$= \frac{\Gamma(r+k+1)}{\Gamma(k+1)\Gamma(r)} p^{r-1} (1-p)^{k}$$

$$= \frac{p^{r-1}(1-p)^{k}}{B(r,k+1)}$$

$$\therefore G'(t) = \frac{1}{B(r,k+1)} t^{r-1} (1-t)^{k}$$

$$\therefore \int_0^p G'(t)\,dt = \frac{1}{B(r,k+1)} \int_0^p t^{r-1} (1-t)^{k}\,dt = I_p(r,k+1) = G(p)$$

$$\therefore P[X \le k] = I_p(r,k+1).$$

$\therefore$ The distribution function is an increasing function of p and decreasing function of q.

$$\therefore P[X > k] = 1 - P[X \le k] = 1 - I_p(r,k+1) = I_q(k+1,r)$$

☆ Problem 1. Let X_1 and X_2 be two such random variables such that $X_1 \sim NB(r, p_1)$ and $X_2 \sim NB(r, p_2)$ and $p_1 > p_2$. Show that for any non-negative integer k, $P[X_1 \le k] > P[X_2 \le k]$.

Proof: $P[X_1 \le k] = I_{p_1}(r, k+1)$ & $P[X_2 \le k] = I_{p_2}(r, k+1)$

$$I_{p_1}(r, k+1) = \int_0^{p_1} \frac{t^{r-1}(1-t)^k \, dt}{B(r, k+1)} = \int_0^{p_2} \frac{t^{r-1}(1-t)^k \, dt}{B(r, k+1)} + \int_{p_2}^{p_1} \frac{t^{r-1}(1-t)^k \, dt}{B(r, k+1)}$$

$$> \int_0^{p_2} \frac{t^{r-1}(1-t)^k \, dt}{B(r, k+1)} \quad \left[\text{since the other term is positive}\right]$$

$$\therefore I_{p_1}(r, k+1) > I_{p_2}(r, k+1)$$

$$\Rightarrow P[X_1 \le k] > P[X_2 \le k]$$

Relationship with Binomial Distribution :—

Result: If $X \sim NB(r, p)$ then for any non-negative integer k,

i) $P[X \le k] = P[Y \ge r]$ where, $Y \sim Bin(r+k, p)$

ii) $P[X \le k] = P[Z \le k]$ where, $Z \sim Bin(r+k, q)$

Proof:—

$$P[X \le k] = I_p(r, k+1)$$

if $Z \sim Bin(r+k, q)$

$$P[Z \le k] = I_p(r, k+1)$$

Now, $P[Z \le k] = P[r+k-Z \ge r]$, putting $r+k-Z = Y$

$= P[Y \ge r]$ where $Y \sim Bin(r+k, p)$

Since X denotes the no. of failures preceeding the rth success.

$(r+X)$ denotes the no. of trials required to get r successes.

Now,

$[X \le k] = [r+X \le r+k]$ = [The number of trials required to get r success is less than equal to $r+k$]

$\Leftrightarrow$ [the number of success in $r+k$ trials is greater than equal to r]

if Y denotes the number of success in $(r+k)$ trials,

$$[X \le k] \Leftrightarrow [Y \ge r]$$

$\therefore P[X \le k] = P[Y \ge r]$, where $Y \sim Bin(r+k, p)$

Again, $P[Y \ge r] = P[r+k-Y \le k]$

Let, $Z = r+k-Y$, $\therefore P[Z \le k] = P[Y \ge r] = P[X \le k]$

where $Z \sim Bin(r+k, q)$.

Relationship Between Negative Binomial and Binomial Distribution:-

24) X = Number of failures preceeding rth success in a sequence of Bernoulli trials with success probability 'p'.

$$P[X \le k] = P[X+r \le k+r]$$
$$= P[N \le k+r]$$
$$= P[Z \ge r]$$

$N = X+r$ = Number of trials required to get r successes.

Z = Number of success out of $(k+r)$ trials.

and $Z \sim Bin(k+r, p)$.

Recursion Relation for Central Moments:-

If $X \sim NB(r, p)$, then

$$\mu_k = E(X-E(X))^k$$
$$= \sum_{x=0}^{\infty}\left[x - \frac{r(1-p)}{p}\right]^k \binom{r+x-1}{x} p^r (1-p)^x \quad , q = 1-p.$$

Differentiating both sides w.r.t. p, we get —

$$\frac{d\mu_k}{dp} = \sum_x \binom{r+x-1}{x}\Bigg[(-rk)\left(x - \frac{r(1-p)}{p}\right)^{k-1} p^r (1-p)^x \left(\frac{-p+p-1}{p^2}\right)$$
$$+ \left(x - \frac{r(1-p)}{p}\right)^k r p^{r-1}(1-p)^x - \left(x - \frac{rq}{p}\right)^k p^r x (1-p)^{x-1}\Bigg]$$

$$= \sum_x \binom{x+r-1}{x}\frac{rk}{p^2}\left(x - \frac{rq}{p}\right)^{k-1} p^r q^x$$
$$+ \sum_x \binom{r+x-1}{x}\left(x - \frac{rq}{p}\right)^k p^{r-1}(1-p)^{x-1}\{r(1-p) - xp\}$$

$$= \frac{rk}{p^2}\mu_{k-1} + \sum_x \binom{r+x-1}{x}\left(x - \frac{rq}{p}\right)^{k+1} p^{r-1}(1-p)^{x-1}.(-p)$$

$$= \frac{rk}{p^2}\mu_{k-1} - \frac{1}{q}\mu_{k+1}$$

$$\therefore \boxed{\mu_{k+1} = q\left[\frac{rk}{p^2}\mu_{k-1} - \frac{d}{dp}\mu_k\right]}$$

Put, $k=1$, $\mu_0 = 1, \mu_1 = 0$,

$$\mu_2 = q\left[\frac{r.1}{p^2} - 0\right] = \frac{rq}{p^2}$$

Putting $k=2$,

$$\mu_3 = q\left[0 + \frac{(1+q)r}{p^3}\right] = \frac{rq(1+q)}{p^3}$$

putting $k=3$,

$$\mu_4 = q\left[\frac{rk}{p^2}\cdot\frac{rq}{p^2} - \frac{d}{dp}\cdot\frac{rq(1+q)}{p^3}\right]$$

$$= q\left[\frac{3r^2q}{p^4} - r\cdot\frac{p^3(2p-3)-(1-p)(2-p)3p^2}{p^6}\right]$$

$$= q\left[\frac{3r^2q}{p^4} - r\cdot\frac{p(2p-3)-q(1+q)3}{p^4}\right]$$

$$= q\left[\frac{3r^2q}{p^4} - r\cdot\frac{2p^2-3p-3q-3q^2}{p^4}\right]$$

$$= q\left[\frac{3r^2q}{p^4} - \frac{(2p^2-3q^2-3)r}{p^4}\right]$$

$$= q\left[\frac{3r^2q+3rq^2-2r(1-q)^2+3r}{p^4}\right]$$

$$= q\left[\frac{3r^2q+3rq^2-2r(1-2q+q^2)+3r}{p^4}\right]$$

$$= \frac{3r^2q^2+rq^3+4rq^2+rq}{p^4}$$

$$\therefore\ \gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{(rq)(1+q)/p^3}{(rq/p^2)^{3/2}} = \frac{1+q}{\sqrt{rq}}$$

$$\therefore\ \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{(3r^2q^2+rq^3+4rq^2+rq)/p^4}{[rq/p^2]^2} - 3$$

$$= \left(3+q+\frac{4}{r}+\frac{1}{rq}\right)-3$$

$$= \left(q+\frac{4}{r}+\frac{1}{rq}\right).$$

Raw Moment :- k-th order raw moment of X is, —

$$\mu_k' = E[X^k] = \sum_{x=0}^{\infty} x^k f(x)$$

$$= \sum_{x=0}^{\infty} x^{k-1}\cdot x\cdot\frac{(r+x-1)!}{x!\,(r-1)!}p^r q^x$$

$$= \frac{rq}{p}\sum_{x=1}^{\infty} x^{k-1}\frac{(r+x-1)!}{(x-1)!\,r!}p^{r+1}q^{x-1}$$

$$= \frac{rq}{p}\sum_{x'=0}^{\infty}(1+x')^{k-1}\binom{r+x'}{x'}p^{r+1}q^{x'} \quad [\text{put, } x-1=x']$$

$$= \frac{rq}{p}E(1+Y)^{k-1}, \text{ where } Y\sim NB(r+1,p)$$

Mean Deviation About Mean:

Let $X \sim NB(r, p)$

$$E(X) = \mu = \frac{rq}{p}.$$

$$MD_\mu(X) = E|X-\mu| = \sum_{x=0}^{\infty} \left|x - \frac{rq}{p}\right| f(x)$$

$$= 2\sum_{x=0}^{[\mu]} \left(\frac{rq}{p} - x\right) f(x)$$

$$= \frac{2}{p}\sum_{x=0}^{[\mu]} \left(rq - (1-q)x\right) f(x)$$

$$= \frac{2}{p}\sum_{x=0}^{[\mu]} \left[(x+r)q - x\right] f(x)$$

$$= \frac{2}{p}\sum_{x=0}^{[\mu]} (x+r)q f(x) - \frac{2}{p}\sum_{x=0}^{[\mu]} x f(x)$$

Let, $g(x) = x f(x) = x \cdot \frac{(r+x-1)!}{(r-1)!\,x!} p^r q^x = q \frac{(r+x-1)!}{(x-1)!\,(r-1)!} p^r q^{x-1}$

$\therefore g(x+1) = q \cdot \frac{(r+x)!}{x!\,(r-1)!} p^r q^x = q\cdot(x+r) f(x)$

$$MD_\mu(X) = \frac{2}{p}\left[\sum_{x=0}^{[\mu]} \{g(x+1) - g(x)\}\right]$$

$$= \frac{2}{p} g([\mu]+1)$$

$$= \frac{2}{p}\left[([\mu]+1) f([\mu]+1)\right]$$

$$= \frac{2}{p}\mu_0 \; P[X = \mu_0]$$

[Let $[\mu]+1 = \mu_0$ $\Rightarrow [\frac{rq}{p}]+1 = \mu_0$]

Now, $P[X = \mu_0] = \binom{\mu_0 + r - 1}{\mu_0} p^r q^{\mu_0}$

$$= \frac{(\mu_0 + r - 1)!}{\mu_0!\,(r-1)!} p^r q^{\mu_0}$$

$n! \simeq \sqrt{2\pi}\, e^{-n} . n^{n+\frac{1}{2}}$ for large n, by Stirling's approximation.

Using this we get —

$$MD_\mu(X) = \frac{1}{2}\sqrt{\frac{2rq}{\pi}}$$

☆ 2. For a sequence of Bernoulli trials, let X be the number of trials required to get 'n' successes where n is a fixed integer, if 'p' be the probability of success for a single trial, then show that,

$$E\left[\frac{n}{X}\right] = \sum_{i=1}^{n-1} (-1)^{i-1} \frac{n}{n-i} \left(\frac{p}{q}\right)^i + \left(-\frac{p}{q}\right)^n n \ln P.$$

Proof:- $f(x) = \binom{x-1}{n-1} p^n q^{x-n}$

Now, $E\left(\frac{n}{X}\right) = \sum_{x=n}^{\infty} \frac{n}{x} \binom{x-1}{n-1} p^n q^{x-n}$

$$= \left(\frac{p}{q}\right)^n n \sum_{x=n}^{\infty} \binom{x-1}{n-1} \frac{q^x}{x}$$

$$= \left(\frac{p}{q}\right)^n \cdot n \sum_{x=n}^{\infty} \binom{x-1}{n-1} \cdot \int_0^q t^{x-1}\, dt$$

$$= \left(\frac{p}{q}\right)^n \cdot n \cdot \sum_{x=n}^{\infty} \binom{x-1}{n-1} \int_0^q t^{n-1} \cdot t^{x-n}\, dt$$

$$= n \left(\frac{p}{q}\right)^n \int_0^q t^{n-1} \sum_{x'=0}^{\infty} \binom{x'+n-1}{x'} t^{x'}\, dt$$

[Let $(x-n) = x'$]

$$= \left(\frac{p}{q}\right)^n \cdot n \int_0^q t^{n-1} (1-t)^{-n}\, dt$$

Let, $I(n-1, -n) = \int_0^q t^{n-1} (1-t)^{-n}\, dt$

$$= \left[\frac{-(1-t)^{-n+1}}{(1-n)} \cdot t^{n-1}\right]_0^q - \int_0^q \frac{(\cancel{n-1}) t^{n-2} \cdot (1-t)^{-(n-1)}}{(\cancel{n-1})}\, dt$$

$$= \frac{(1-q)^{-(n-1)} \cdot q^{n-1}}{(n-1)} - I\left[(n-2), -(n-1)\right]$$

$$= \frac{\left(\frac{q}{p}\right)^{n-1}}{(n-1)} - I\left[(n-2), -(n-1)\right]$$

$$I\left[(n-2), -(n-1)\right] = \frac{\left(\frac{q}{p}\right)^{n-2}}{n-2} - I\left[(n-3), -(n-2)\right]$$

$$\therefore I(n-1, -n) = \frac{1}{n-1}\left(\frac{q}{p}\right)^{n-1} - \frac{1}{n-2}\left(\frac{q}{p}\right)^{n-2} + \left(\frac{1}{n-3} \cdot \left(\frac{q}{p}\right)^{n-3} \cdots\right.$$

$$\cdots + (-1)^{n-2} \frac{1}{n-(n-1)} \left(\frac{q}{p}\right) + (-1)^n \ln P$$

$$\left[\because I(0, -1) = \int_0^q (1-t)^{-1} dt = \left[-\ln(1-t)\right]_0^q = -\ln P\right]$$

$$\therefore I(n-1,-n) = \sum_{i=1}^{n-1}\left(\frac{p}{q}\right)^{n-i}\frac{(-1)^{i-1}}{n-i} + (-1)^n \ln p \qquad \cdot(28)$$

$$\therefore E\left[\frac{n}{X}\right] = \sum_{i=1}^{n-1}(-1)^{i-1}\frac{n}{n-i}\left(\frac{p}{q}\right)^{i} + \left(-\frac{p}{q}\right)^{n} n.\ln p \quad \text{[Proved]}$$

☆ 3. Obtain the probability that in a sequence of Bernoulli trials with p for 'a' successes will occur before 'b' failures. Hence obtain the identity

$$\sum_{k=0}^{b-1}\binom{a+b-1}{a+k}p^k q^{b-k-1} = \sum_{k=0}^{b-1}\binom{a+k-1}{k}q^k$$

Solⁿ. Let, X denotes the number of failures preceeding the a th success, —

$$X \sim NB(a,p)$$

$\therefore$ Required probability :— $P[X \le b-1]$

$$= \sum_{x=0}^{b-1}\binom{a+x-1}{x}p^a q^x$$

Let, Y denotes the number of failures in $\overline{a+b-1}$ trials, —

$$Y \sim Bin(a+b-1, q)$$

$\therefore$ Required probability :— $P[X \le b-1]$

$$= \sum_{x=0}^{b-1}\binom{a+b-1}{x}q^x p^{a+b-1-x}$$

$$\therefore \sum_{x=0}^{b-1}\binom{a+x-1}{x}p^a q^x = \sum_{x=0}^{b-1}\binom{a+b-1}{x}q^x p^a . p^{b-1-x}$$

Let, in R.H.S $b-1-x=j$

$$\Rightarrow \sum_{x=0}^{b-1}\binom{a+x-1}{x}q^x = \sum_{j=0}^{b-1}\binom{a+b-1}{j}q^{b-1-j}p^j$$

$$\therefore \sum_{k=0}^{b-1}\binom{a+k-1}{k}q^k = \sum_{k=0}^{b-1}\binom{a+b-1}{k}q^{b-1-k}.p^k$$

[Proved]

Scanned by CamScanner

Mode of a Negative Binomial Distribution : —

Let $X \sim NB(r, p)$, the p.m.f. of X is $f(x) = \binom{x+r-1}{x} p^r q^x$

The mode of the distribution is the value(s) of x for which $f(x)$ is maximum. for $x = 1, 2, \ldots \infty$

Let us consider the ratio, —

$$\frac{f(x)}{f(x-1)} = \frac{\binom{x+r-1}{x} p^r q^x}{\binom{x+r-2}{x} p^r q^{x-1}} = \frac{r+x-1}{x} \cdot q$$

$$\therefore \frac{f(x)}{f(x-1)} \gtreqless 1 \text{ according as } (r+x-1)(1-p) \gtreqless x$$

$$\Rightarrow (r-1)(1-p) \gtreqless xp$$

$$\Rightarrow x \lesseqgtr (r-1)\frac{(1-p)}{p}$$

$\therefore f(x) \gtreqless f(x-1)$ according as $x \lesseqgtr (r-1)\frac{(1-p)}{p}$.

Case-I $(r-1)\frac{(1-p)}{p}$ is an integer $= k$ (say)

Now, $x = k$

$\Rightarrow f(k) = f(k-1) \cdots\cdots$ ①

for, $x < k$

$\Rightarrow f(x) > f(x-1)$ for $x = 1, \ldots, k-1$.

$\Rightarrow f(0) < f(1) < \cdots < f(k-1) \cdots\cdots$ ②

for, $x > k$

$\Rightarrow f(x) < f(x-1)$ for $x = k+1, k+2, \ldots \infty$

$\Rightarrow f(k) > f(k+1) > \cdots > \infty \cdots\cdots$ ③

Combining ①, ② & ③, we get —

$$f(0) < f(1) < \cdots < f(k-1) = f(k) > f(k+1) > \cdots$$

$\therefore f(x)$ is maximum at $x = k$ and $x = k-1$.

$\therefore$ The mode of the distribution are at

$$(r-1)\frac{(1-p)}{p} \text{ and } (r-1)\cdot\frac{(1-p)}{p} - 1.$$

Case-II If $(r-1)\frac{(1-p)}{p}$ is not an integer.

Let, $k_0 = \left[(r-1)\frac{(1-p)}{p}\right]$ = largest integer obtained in $\frac{(r-1)(1-p)}{p}$

Obviously,

$$k_0 < (r-1)\cdot\frac{(1-p)}{p} < k_0+1$$

Now, $x = k_0$,

$$\Rightarrow x < (r-1)\cdot\frac{(1-p)}{p}$$

$$\Rightarrow f(k_0) > f(k_0-1) \;\cdots\cdots (4)$$

$$x < k_0,$$

$$\Rightarrow x < (r-1)\frac{(1-p)}{p}$$

$$\Rightarrow f(x) > f(x-1) \quad \text{for } x = 1, 2, \ldots, k_0-1$$

$$\Rightarrow f(0) < f(1) < \cdots < f(k_0-1) \;\cdots\cdots (5)$$

$$x > k_0$$

$$\Rightarrow x \geq k_0+1 > (r-1)\cdot\frac{(1-p)}{p}$$

$$\Rightarrow f(x) < f(x-1) \quad \text{for } x = k_0+1, k_0+2, \ldots \infty$$

$$\Rightarrow f(k_0) > f(k_0+1) > \cdots\cdots > \cdots \infty \;\cdots\cdots (6)$$

Combining (4), (5) & (6), we get —

$$f(0) < f(1) < \cdots < f(k_0-1) < f(k_0) > f(k_0+1) > \cdots\cdots$$

$\therefore f(x)$ is maximum at $x = k_0$.

Factorial moment generating function: —

Let $X \sim NB(r, p)$

PGF of X is given by, —

$$P_X(t) = E(t^X) \quad, \text{since } |t| < 1.$$

$$= \sum_{x=0}^{\infty} \binom{x+r-1}{x} p^r q^x t^x$$

$$= p^r (1-qt)^{-r} \quad \left[\because |t| < \frac{1}{q}\right]$$

Now, factorial moment generating function is

$$P_X(1+t) = p^r (1-q-qt)^{-r}$$

$$= \left(1 - \frac{q}{p}t\right)^{-r} = \sum_{x=0}^{\infty} \binom{x+r-1}{x} \left(\frac{qt}{p}\right)^x$$

$$= \sum_{x=0}^{\infty} \binom{x+r-1}{x} \left(\frac{q}{p}\right)^x \cdot x! \cdot \frac{t^x}{x!}$$

$\therefore \mu'_{[x]}$ = coefficient of $\frac{t^x}{x!}$ in the expansion of $P_X(1+t)$

$$= (x+r-1)_x \left(\frac{q}{p}\right)^x$$

(Factorial moment ⟶ Rawmoment ⟶ central moment)

$$\mu_1' = E(X) = \frac{rq}{p} = \text{mean}$$

$$\mu_2' = \mu'_{[2]} + \mu'_{[1]}$$

$$= r(r+1)\frac{q^2}{p^2} + \frac{rq}{p}$$

$$\mu_3' = \mu'_{[3]} + 3\mu'_{[2]} + \mu'_{[1]}$$

$$= r(r+1)(r+2)\frac{q^3}{p^3} + 3.r(r+1)\frac{q^2}{p^2} + \frac{rq}{p}$$

$$\mu_4' = \mu'_{[4]} + 6\mu'_{[3]} + 7\mu'_{[2]} + \mu'_{[1]}$$

$$= r(r+1)(r+2)(r+3)\left(\frac{q}{p}\right)^4 + 6r(r+1)(r+2)\left(\frac{q}{p}\right)^3 + 7.r(r+1)\left(\frac{q}{p}\right)^2 + \frac{rq}{p}$$

☆ 4. Suppose the probability that a workman chosen at random will possess accident proneness of intensity λ to $\lambda + d\lambda$ and that he produces x accidents is

$$f(\lambda)\,d\lambda\, p(x,\lambda),$$

where $f(\lambda) = \frac{\gamma^\alpha}{\Gamma(\alpha)} \exp[-\gamma\lambda]\,\lambda^{\alpha-1}, \quad 0<\lambda<\infty,$

and $p(x,\lambda) = \frac{\exp[-\lambda]\,\lambda^x}{x!}, \quad x = 0,1,2,\ldots$

Show that the probability that a workman chosen at random (with unknown accident proneness) will produce x accidents follows a negative Binomial distn.,

where $p = \frac{\gamma}{(\gamma+1)}$ and $r = \alpha$.

Solution: —

$$f(\lambda) = \lim_{d\lambda \to 0} \frac{F(\lambda+d\lambda) - F(\lambda)}{d\lambda}$$

$$= \lim_{d\lambda\to 0} \frac{P(\lambda < X < \lambda + d\lambda)}{d\lambda}, \ldots$$

For small $d\lambda$, (≥ 0)

$P[\lambda < X < \lambda + d\lambda] \simeq f(\lambda)\,d\lambda$ is the probability element.

The p.m.f. of X is,

$$f(x) = \int_0^\infty p(x,\lambda) f(\lambda)\, d\lambda$$

$$= \int_0^\infty \frac{e^{-\lambda}.\lambda^x}{x!} \cdot \frac{\gamma^\alpha}{\Gamma(\alpha)} \cdot e^{-\gamma\lambda} . \lambda^{\alpha-1}$$

$$= \frac{\gamma^\alpha}{x!\,\Gamma(\alpha)} \int_0^\infty e^{-(1+\gamma)\lambda} . \lambda^{x+\alpha-1}\, d\lambda$$

$$= \frac{\gamma^\alpha}{x!\,\Gamma(\alpha)} \int_0^\infty \frac{e^{-(1+\gamma)\lambda} \{(1+\gamma)\lambda\}^{x+\alpha-1}\, d\lambda}{(1+\gamma)^{x+\alpha}}$$

$$= \frac{\gamma^\alpha}{x!\,\Gamma(\alpha)} \cdot \Gamma(x+\alpha) \cdot \frac{1}{(1+\gamma)^{x+\alpha}} \quad , x = 0,1,2,\ldots$$

$$= \frac{(x+\alpha-1)!}{x!\,(\alpha-1)!} \cdot \gamma^\alpha \cdot \frac{1}{(1+\gamma)^\alpha} \cdot \frac{1}{(1+\gamma)^x}$$

$$= \binom{x+\alpha-1}{x} \left(\frac{\gamma}{1+\gamma}\right)^\alpha \cdot \left(\frac{1}{1+\gamma}\right)^x$$

Here $X \sim NB(r,p)$, where $r = \alpha$, $p = \left(\frac{\gamma}{1+\gamma}\right)$

☆ 5. <u>Relationship between Binomial & Negative Binomial distribution.</u>

Let $X \sim Bin(n,p)$ and $Y \sim NB(r,p)$ denotes the no. of trials required to get rth success in a sequence of independent Bernoulli trials. Then S.T. $F_X(r-1) = 1 - F_Y(n)$. ///

<u>Soln</u> Consider an event,

" there are r or more successes in the first n bernoullian trials " which is equivalent to the event, " required to get the rth success ".

Hence, $\{X \geq r\} = \{Y \leq r\}$

$$\Rightarrow P[X \geq r] = P[Y \leq n]$$

$$\Rightarrow 1 - P[X \leq r-1] = P[Y \leq n]$$

$$\Rightarrow 1 - F_X(r-1) = F_Y(n)$$

<u>Remark:</u>— In $Bin(n,p)$ distribution, we count the no. of successes in 'n' independent Bernoulli trials where as in $NB(r,p)$, $r \in \mathbb{N}$, we count the no. of trials required to get the rth success. Therefore, <u>Negative Binomial is also inverse Binomial Sampling.</u>

☆6. If $X \sim NB(r, p)$ and $Z = \dfrac{X - \frac{rq}{p}}{\frac{\sqrt{rq}}{p}}$, then $M_Z(t) \to e^{t^2/2}$ for large r.

<u>Ans:-</u>

$$M_Z(t) = E(e^{tZ})$$
$$= E\left[e^{t\left(\frac{X - \frac{rq}{p}}{\sqrt{rq}/p}\right)}\right]$$
$$= e^{-t\sqrt{rq}}\, M_X\left(\frac{pt}{\sqrt{rq}}\right)$$

We know, $M_X(t) = p^r(1 - qe^t)^{-r}$

$$M_Z(t) = e^{-t\sqrt{rq}}.p^r\left(1 - qe^{pt/\sqrt{rq}}\right)^{-r}$$
$$= \left[\frac{1}{p}e^{t\sqrt{q/r}} - \frac{q}{p}e^{\frac{pt}{\sqrt{rq}} + t\sqrt{\frac{q}{r}}}\right]^{-r}$$
$$= \left[\frac{1}{p}e^{t\sqrt{\frac{q}{r}}} - \frac{q}{p}e^{t/\sqrt{rq}}\right]^{-r}$$

$$= \left[\frac{1}{p}\left(1 + t\sqrt{\frac{q}{r}} + t^2 \cdot \frac{q}{2r} + O\left(r^{-3/2}\right)\right) - \frac{q}{p}\left[1 + \frac{t}{\sqrt{rq}} + \frac{t^2}{2rq} + O\left(r^{-3/2}\right)\right]\right]^{-r}$$

$$= \left[1 + \frac{t^2}{2r} + O\left(r^{-3/2}\right)\right]^{-r}$$

$$\therefore \lim_{r\to\infty} M_Z(t) = \lim_{r\to\infty}\left[1 + \frac{t^2}{2r}\right]^{-r}$$
$$= e^{t^2/2}$$

POISSON DISTRIBUTION :

A discrete random variable X with mass points $0,1,2,\ldots\infty$ is said to follow Poisson distribution with parameter $\lambda(>0)$ if its p.m.f. is given by, —

$$f(x) = \begin{cases} \dfrac{e^{-\lambda}.\lambda^{x}}{x!} & , x = 0,1,2,\ldots \\ & \lambda > 0 \\ 0 & \text{ow} \end{cases}$$

We denote $X \sim P(\lambda)$.

<u>Example:</u> → Some real life situations where the Poisson law is appropriate, —

i) Distribution of number of misprints on a certain page of an encyclopaedia.
ii) No. of cross connections during an hour in a busy telephone exchange.
iii) No. of total road accidents during hour in a five point junction.
iv) No. of defects in a piece of cloth of specified length.
v) No. of defective items in a factory.

<u>Uses:</u> — Poisson Distribution is used to model the probability of occurance of rare events.

<u>Remark:</u> — The number of misprints on a page of a book will be — Here the probability that each letter typed on a page is misprinted is small and the no. of letters on a page is quite large. But, in the 1st proof of a book, the probability that a letter typed is misprinted is not small enough and consequently "the no. of misprints on a page of a book" will not follow approximately a Poisson Distribution.

In practice, great care has to be taken to avoid erroneously application to Poisson distribution for counting variables. For e.g., in studying the distribution of deaths of insects form a clusters which is inconsistent with the assumption of Poisson Process.

Poisson Process : Suppose that we are observing the occurance of an event in time, a quantity $\mu(>0)$ such that :

i) The probability that exactly one happening will occur in a small time interval of length 't' is μt.

ii) The probability that more than one happening in a time interval of length 't' is negligible.

iii) The no. of happening in non-overlapping intervals are independent.

Then the no. of occurances in the interval of the length 'T' follows a "Poisson distribution" with mean $= \mu T$. Here μ is known as the mean rate of occurance. Then "the no. of occurances in the time interval of length $\frac{T}{2}$" follows "Poisson Distribution" with mean $= \frac{\mu T}{2}$.

From the conditions of "Poisson process" certain random experiments revolving counts of occurances of an event in time, space or length can be realistically modelled by Poisson Distribution.

Result :- Let $X \sim Bin(n, p)$, suppose the following conditiom are satisfied,

i) $n \to \infty$ (i.e. the no. of trials is very large).

ii) $p \to 0$ (i.e. the probability of success is very small)

iii) $np = \lambda$ (finite) [i.e. the average number of trial is finite].

Then the distn. of X will converge to a Poisson distn. with parameter λ.

Proof:- for some fixed small x,

$$P[X = x] = \binom{n}{x} p^x q^{n-x} \quad , x = 0(1)n.$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \quad [\because np=\lambda]$$

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

As $n \to \infty$, $p \to 0$, $np = \lambda$,

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right) \to 1,$$

$$\left(1-\frac{\lambda}{n}\right)^x \to 1, \text{ and } \left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

$$\therefore f(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} = \lim_{n\to\infty} P[X = x]$$

$\therefore$ limiting case of Bin. distn. when $n \to \infty$, $p \to 0$ but $np = \lambda$ (finite) can be approximated by $\underset{\sim}{Pois}(\lambda)$.

Another method: $X \sim Bin(n, p)$

$$M_X(t) = (q + pe^t)^n$$
$$= \{1 - p(1-e^t)\}^n$$
$$= \left\{1 - \frac{np}{n}(1-e^t)\right\}^n$$
$$= \left\{1 - \frac{\lambda}{n}(1-e^t)\right\}^n$$

$$\therefore \lim_{n\to\infty} M_X(t) = \exp[-\lambda(1-e^t)]$$
$$= e^{\lambda(e^t - 1)}$$

which is the m.g.f. of a Poisson distn. with parameter λ.

Hence, by uniqueness of m.g.f. as $n \to \infty$, $p \to 0$, $np = \lambda$,
X is a Poisson R.V.

Result: — Let $X \sim NB(r, p)$. Suppose the following conditions are satisfied,

i) $r \to \infty$
ii) $q \to 0$
iii) $rq = \lambda$ (finite)

Then the distn. of X will converge to a Poisson distn. with parameter λ.

Proof: — for some fixed x,

$$P[X = x] = \binom{x+r-1}{x} p^r q^x, \quad x = 0, 1, 2, \ldots$$

$$= \frac{(x+r-1)\cdots\cdots r}{x!}\left(1-\frac{\lambda}{r}\right)^r\left(\frac{\lambda}{r}\right)^x$$

$$= \frac{\left(\frac{x-1}{r}+1\right)\left(\frac{x-2}{r}+1\right)\cdots\cdots 1}{x!} \cdot \lambda^x \left(1-\frac{\lambda}{r}\right)^r$$

$$\because r \to \infty, \quad \left(\frac{x-1}{r}+1\right)\left(\frac{x-2}{r}+1\right)\cdots 1 \to 1$$

$$\& \quad \left(1 - \frac{\lambda}{r}\right)^r \to e^{-\lambda}$$

$$\therefore f(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

Another method: Let $X \sim NB(r, p)$

Suppose, $r \uparrow \infty$, $q \downarrow 0$ and $q = 1-p \ni rq = \lambda$ (finite)

P.G.F. of X is —

$$P_X(t) = p^r(1-qt)^{-r}$$

$$= (1-q)^r \left(1-\frac{qt}{}\right)^{-r}$$

$$= \left(1-\frac{\lambda}{r}\right)^r \left(1-\frac{\lambda t}{r}\right)^{-r}$$

$$\rightarrow e^{-\lambda}.e^{\lambda t} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda t)^x}{x!}$$

$$\therefore P[X=x] = \frac{e^{-\lambda}.\lambda^x}{x!}$$

[Poisson Distn. as a limiting case of negative Binomial Distribution]

Recurrence Relation Regarding Probability: —

If $X \sim P(\lambda)$, then, —

$$\frac{P(X=x)}{P(X=x-1)} = \frac{e^{-\lambda}.\frac{\lambda^x}{x!}}{e^{-\lambda}.\frac{\lambda^{x-1}}{(x-1)!}} = \frac{\lambda}{x}$$

$$\therefore P(X=x) = \frac{\lambda}{x}.P(X=x-1), x=1,2,3,\ldots$$

Factorial Moments: — Let $X \sim Pois(\lambda)$, then the rth order factorial moment of X is given by —

$$\mu_{[r]} = E[(X)_r] = E[X(X-1)\cdots(X-r+1)]$$

$$= \sum_{x=0}^{\infty}(x)_r f(x)$$

$$= \sum_{x=0}^{\infty}(x)_r . \frac{e^{-\lambda}.\lambda^x}{x!}$$

$$= \sum_{x=r}^{\infty}\frac{e^{-\lambda}.\lambda^x}{(x-r)!} \quad [\because x! = (x)_r (x-r)!$$

$$= e^{-\lambda}\lambda^r\sum_{x=r}^{\infty}\frac{\lambda^{x-r}}{(x-r)!}$$

$$= e^{-\lambda}.\lambda^r.e^{\lambda}$$

$$= \lambda^r = \mu'_{[r]}$$

Now, we can get raw moments from factorial moments and central moments from raw moments.

$\mu_1' = \mu'_{[1]} = \lambda$

$\mu_2' = \mu'_{[2]} + \mu'_{[1]}$

$= \lambda^2 + \lambda$

$\mu_3' = \mu'_{[3]} + 3\mu'_{[2]} + \mu'_{[1]}$

$= \lambda^3 + 3\lambda^2 + \lambda$

$\mu_4' = \mu'_{[4]} + 6\mu'_{[3]} + 7\mu'_{[2]} + \mu'_{[1]}$

$= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$

$\therefore E(X) = \mu'_{[1]} = \lambda$

& $Var(X) = \mu'_{[2]} + \mu'_{[1]} - [\mu'_{[1]}]^2$ $\boxed{\therefore \text{ Mean} = \text{Variance}}$

$= \lambda^2 + \lambda - \lambda^2 = \lambda$

Central moments:–

$\mu'_{[1]} = \mu_1' = \mu_1 = \lambda$

$\mu_2 = \lambda$

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

$= \lambda^3 + 3\lambda^2 + \lambda - 3(\lambda^2 + \lambda)\lambda + 2\lambda^3$

$= \lambda$.

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

$= 3\lambda^2 + \lambda$.

Skewness:–
$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda}$$

$\therefore \gamma_1 = \frac{1}{\sqrt{\lambda}}$, since $\lambda > 0$, Thus the Poisson distribution is positively skewed.

Kurtosis:–
$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2 + \lambda}{\lambda^2} = 3 + \frac{1}{\lambda}$$

$\therefore \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda} > 0$

Thus the distribution is leptokurtic.

Probability Generating Function: $\rightarrow$

$$PGF = P(t) = E(t^X) = \sum_{x=0}^{\infty} t^x \cdot \frac{e^{-\lambda}.\lambda^x}{x!}$$

$$= e^{\lambda(t-1)}$$

Factorial moment Generating function: $\longrightarrow$

$$P(1+t) = e^{\lambda t}$$

$$= \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!}$$

coefficient of $\frac{t^x}{x!} = \lambda^x = \mu'_{[x]}$

Moment Generating Function: $\longrightarrow$

$$MGF = M(t) = e^{\lambda(e^t-1)} \quad \forall\, t \in \mathbb{R}$$

$$= e^{\lambda\left(\sum_{x=1}^{\infty} \frac{t^x}{x!}\right)}$$

$$= 1 + \lambda \sum_{x=1}^{\infty} \frac{t^x}{x!} + \frac{\lambda^2}{2}\left(\sum_{x=1}^{\infty} \frac{t^x}{x!}\right)^2 + \cdots\cdots \infty$$

$$= 1 + \lambda\left(t + \frac{t^2}{2} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots\cdots \infty\right)$$

$$+ \frac{\lambda^2}{2}\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots\cdots \infty\right)^2$$

$$+ \frac{\lambda^3}{3!}\left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots \infty\right)^3$$

$$+ \frac{\lambda^4}{4!}\left(t + \frac{t^2}{2!} + \cdots \infty\right)^4 + \cdots\cdots \infty$$

$\therefore \mu'_1 = \lambda$; $\mu'_2 = \lambda + \lambda^2$; $\mu'_3 = \lambda^3 + 3\lambda^2 + \lambda$;

$\mu'_4 = \lambda^4 + 7\lambda^2 + 6\lambda^3 + \lambda$

Central moment Generating Function: $\longrightarrow$

$$M_*(t) = E\left[e^{t(X-\lambda)}\right]$$

$$= e^{-\lambda t} E(e^{tX})$$

$$= e^{-\lambda t} . e^{\lambda(e^t-1)}$$

$$= e^{\lambda(e^t-1-t)} = e^{\lambda\left(\frac{t^2}{2!} + \frac{t^3}{3!} + \cdots \infty\right)}$$

$$= 1 + \lambda\left(\frac{t^2}{2!} + \frac{t^3}{3!} + \cdots \infty\right) + \frac{\lambda^2}{2}\left(\frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\cdots \infty\right)$$

$$+ \cdots\cdots \infty$$

$\therefore \mu_2 = \lambda$, $\mu_3 = \lambda$, $\mu_4 = \lambda + 3\lambda^2$.

Cumulant Generating Function :

$$K(t) = CGF = \ln M(t)$$

$$= \ln e^{\lambda(e^t - 1)} = \lambda(e^t - 1) = \sum_{x=1}^{\infty} \lambda \cdot \frac{t^x}{x!}$$

$$\Rightarrow k_i = \lambda \;\; \forall \; i = 1, 2, \ldots, \infty.$$

* All the cumulants of Poisson Distribution are equal. *

$$\therefore \; \mu_1' = k_1 = \lambda$$

$$\mu_2 = k_2 = \lambda$$

$$\mu_3 = k_3 = \lambda$$

$$\mu_4 = k_4 + 3k_2^2$$

$$= \lambda + 3\lambda^2.$$

Recursion Relation for Central moments :

$$\mu_r = E(X-\lambda)^r$$

$$= \sum_{x=0}^{\infty} (x-\lambda)^r \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

Now, differentiating both sides w.r.t. λ, we get, —

$$\frac{d\mu_r}{d\lambda} = \sum_{x=0}^{\infty} (-r)(x-\lambda)^{r-1} \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} + \sum_{x=0}^{\infty} (x-\lambda)^r \cdot \frac{1}{x!} \left[-e^{-\lambda} \cdot \lambda^x + x e^{-\lambda} \cdot \lambda^{x-1} \right]$$

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

$$= -r\mu_{r-1} + \frac{1}{\lambda} \mu_{r+1}$$

$$\therefore \boxed{\mu_{r+1} = \lambda \left[\frac{d}{d\lambda} \mu_r + r\mu_{r-1} \right]}, \; r = 1, 2, \ldots$$

$$\therefore \; \mu_1 = 0, \; \mu_0 = 1.$$

Putting, $r = 1$,

$$\mu_2 = \lambda[0+1] = \lambda$$

Putting, $r = 2$,

$$\mu_3 = \lambda[1 + 2 \times 0] = \lambda$$

Putting $r = 3$,

$$\mu_4 = \lambda[1 + 3\lambda]$$

$$= \lambda + 3\lambda^2.$$

Mode of the Poisson Distribution: $\longrightarrow$

$X \sim P(\lambda)$.

$$f(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$\therefore \frac{f(x)}{f(x-1)} = \frac{\lambda}{x}, \quad x = 1, 2, \ldots \infty$$

$$\therefore \frac{f(x)}{f(x-1)} \gtreqless 1 \text{ according as } x \lesseqgtr \lambda.$$

Case-I: λ is an integer.

$f(x-1) < f(x)$ for $x = 1, 2, \ldots, \lambda - 1$

$f(x-1) = f(x)$ for $x = \lambda$

$f(x-1) > f(x)$ for $x = \lambda+1, \lambda+2, \ldots$

$\therefore f(0) < f(1) < f(2) < \ldots < f(\lambda-1) = f(\lambda) > f(\lambda+1) > f(\lambda+2) > \ldots$

Thus, λ and $\lambda - 1$ are the two modes of the distn.

Case-II: λ is not an integer, $[\lambda] = m$,

$\therefore f(x-1) < f(x)$ for $x = 1, 2, \ldots, m$.

$\therefore f(x-1) > f(x)$ for $x = m+1, m+2, \ldots$.

$\therefore f(0) < f(1) < \ldots < f(m) > f(m+1) > \ldots$

$\therefore$ m is the mode of the distn where λ is not an integer, i.e. the largest integer cann't exceed λ.

Mean Deviation about mean: — $X \sim P(\lambda)$, $\mu = \lambda$.

$$\therefore MD_\mu(X) = E|X - \mu|$$

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

$$= 2 \sum_{x=0}^{[\mu]} (\mu - x) \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad \text{Let, } [\mu] = x_0$$

$$= 2 \sum_{x=0}^{x_0} \lambda \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} - 2 \sum_{x=1}^{x_0} \frac{e^{-\lambda} \cdot \lambda^x}{(x-1)!}$$

$$= 2 \sum_{x=0}^{x_0} \frac{e^{-\lambda} \cdot \lambda^{x+1}}{x!} - 2 \sum_{x'=0}^{x_0 - 1} \frac{e^{-\lambda} \cdot \lambda^{x'+1}}{x'!} \qquad [\text{let, } x - 1 = x' \text{ in R.H.S}]$$

$$= 2\lambda \left[\sum_{x=0}^{x_0} \frac{e^{-\lambda} \cdot \lambda^x}{x!} - \sum_{x'=0}^{x_0 - 1} \frac{e^{-\lambda} \cdot \lambda^{x'}}{x'!} \right]$$

$$= 2\lambda \cdot \left[\frac{e^{-\lambda} \cdot \lambda^{x_0}}{x_0!} \right] = \frac{2 \cdot e^{-\lambda} \cdot \lambda^{x_0 + 1}}{x_0!}$$

for large λ, $x_0 \to \lambda$, & $MD_\mu(X) \longrightarrow \dfrac{2e^{-\lambda} \cdot \lambda^{\lambda+1}}{\lambda!}$

Using Stirling's approximation for large λ.

$$MD_\mu(X) \simeq \frac{2 \cdot \cancel{e^{-\lambda}} \cdot \lambda^{\lambda+1}}{\sqrt{2\pi} \cdot \cancel{e^{-\lambda}} \cdot \lambda^{\lambda+\frac{1}{2}}}$$

$$= \sqrt{\frac{2\lambda}{\pi}}$$

Here standard deviation $= \sqrt{\lambda}$

$$\therefore \frac{MD_\mu(X)}{s.d.} = \sqrt{\frac{2\lambda}{\pi}} \cdot \frac{1}{\sqrt{\lambda}} = \sqrt{\frac{2}{\pi}}$$

☆ Problem : 1. If $X \sim P(\lambda = 1)$, then show that $MD_\mu(X)$ is equal to $\left(\frac{2}{e} \times s.d.\right)$.

Soln. :→ $MD_\mu(X) = E|X-\lambda|$, $\lambda = 1$

$$= E|X-1|$$

$$= 2\sum_{x=0}^{1}(1-x) \cdot f(x) \qquad [\because E(X-\lambda)=0]$$

$$= 2 \cdot f(0)$$

$$= 2 \cdot \frac{e^{-1} \cdot 1^0}{0!}$$

$$= \frac{2}{e} \cdot \sqrt{\lambda} \qquad [\text{as } \sqrt{\lambda} = 1]$$

$$= \frac{2}{e} \times s.d. \quad \text{(Proved)}$$

☆ Problem 2. If $X \sim P(\lambda)$. Find $E\left(\frac{1}{1+X}\right)$.

Soln. →

$$E\left(\frac{1}{1+X}\right) = \sum_{j=0}^{\infty} \frac{1}{(1+j)} \cdot e^{-\lambda} \cdot \frac{\lambda^j}{j!}$$

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

$$= \frac{1}{\lambda}\left[\sum_{j'=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^{j'}}{j'!} - e^{-\lambda}\right] \quad ; j' = j+1$$

$$= \frac{1}{\lambda}\left[1 - e^{-\lambda}\right]$$

☆ **Problem 3.** If $X \sim P(\lambda)$ then show that —

(43) $$E[X.g(X)] = \lambda E[g(X+1)]$$

Soln. :→

$$E[Xg(X)] = \sum_{j=1}^{\infty} j g(j) \cdot \frac{e^{-\lambda} \cdot \lambda^j}{j!}$$

$$= \lambda \sum_{j=1}^{\infty} g(j) \cdot \frac{e^{-\lambda} \cdot \lambda^{j-1}}{(j-1)!}$$

$$= \lambda . \sum_{j'=0}^{\infty} g(j'+1) \cdot \frac{e^{-\lambda} \cdot \lambda^{j'}}{j'!} \quad ; j' = j-1$$

$$= \lambda . E[g(X+1)]$$

☆ **Problem 4.** If $X \sim P(\lambda)$, then show that $E(X^r) = \lambda E(X+1)^{r-1}$. Hence find $E(X)$ & $Var(X)$.

Soln. :→ $X \sim P(\lambda)$.

$$E(X^r) = \sum_{j=1}^{\infty} j^r f(j)$$

$$= \sum_{j=1}^{\infty} j^r \cdot \frac{e^{-\lambda} \cdot \lambda^j}{j!}$$

$$= \lambda \sum_{j=1}^{\infty} j^{r-1} \cdot \frac{e^{-\lambda} \cdot \lambda^{j-1}}{(j-1)!}$$

$$= \lambda \sum_{j'=0}^{\infty} (j'+1)^{r-1} \cdot \frac{e^{-\lambda} \cdot \lambda^{j'}}{j'!} \qquad [j' = j-1]$$

$$= \lambda \sum_{j'=0}^{\infty} (j'+1)^{r-1} \cdot f(j')$$

$$= \lambda E(X+1)^{r-1}$$

Now putting, $r=1$, $E(X) = \lambda$

$r=2$, $E(X^2) = \lambda E(X+1)$

$$= \lambda E(X) + \lambda$$

$$= \lambda^2 + \lambda$$

$$\therefore Var(X) = E(X^2) - E^2(X)$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

☆ **Problem 5.** If $X \sim P(\lambda)$ then show that —

$$\mu_{r+1} = \lambda\left[\binom{r}{1}\mu_{r-1} + \binom{r}{2}\mu_{r-2} + \cdots + \binom{r}{r}\mu_0\right]$$

Soln. →

$$\mu_{r+1} = E(X-\lambda)^{r+1}$$
$$= \sum_{j=0}^{\infty}(j-\lambda)^{r+1} f(j)$$
$$= \sum_{j=0}^{\infty}(j-\lambda)^{r}(j-\lambda) f(j)$$
$$= \sum_{j=1}^{\infty}(j-\lambda)^{r}.j.\frac{e^{-\lambda}.\lambda^{j}}{j!} - \lambda\sum_{j=0}^{\infty}(j-\lambda)^{r} f(j)$$
$$= \lambda\sum_{j'=0}^{\infty}(j'+1-\lambda)^{r}.\frac{e^{-\lambda}.\lambda^{j'}}{j'!} - \lambda\mu_r \quad [j'=j-1]$$
$$= \lambda\sum_{j'=0}^{\infty}\left[\sum_{i=0}^{r}\binom{r}{i}(j'-\lambda)^{r-i} f(j')\right] - \lambda\mu_r$$
$$= \lambda\sum_{i=0}^{r}\binom{r}{i}\sum_{j'=0}^{\infty}(j'-\lambda)^{r-i} f(j') - \lambda\mu_r$$
$$= \lambda\sum_{i=0}^{r}\binom{r}{i}\mu_{r-i} - \lambda\mu_r$$
$$= \lambda\sum_{i=1}^{r}\binom{r}{i}\mu_{r-i} \quad \text{(Proved)}$$

☆ **Problem 6.** For the Poisson distribution with parameter λ, S.T.
$E(k.X.e^{-kX}) = \lambda k.e^{\lambda(e^{-k}-1)-k}$, k being a real constant.

Soln. → If $X \sim P(\lambda)$, then —

$$E(k.X.e^{-kX}) = \sum_{x=0}^{\infty} x.k.e^{-kx}.\frac{\lambda^{x}.e^{-\lambda}}{x!}$$
$$= ke^{-\lambda}\sum_{x=0}^{\infty}\frac{x e^{-kx}.\lambda^{x}}{x!}$$
$$= ke^{-\lambda}\sum_{x=1}^{\infty}\frac{(e^{-k}\lambda)^{x}}{(x-1)!}$$
$$= e^{-\lambda}.e^{-k}.\lambda.k\left[\sum_{x=1}^{\infty}\frac{(e^{-k}.\lambda)^{x-1}}{(x-1)!}\right]$$
$$= e^{-(\lambda+k)}.\lambda k\left[\sum_{x'=0}^{\infty}\frac{(e^{-k}.\lambda)^{x'}}{x'!}\right] \quad [\text{taking, } x'=(x-1)]$$
$$= e^{-(\lambda+k)}.\lambda k.e^{\lambda e^{-k}}$$
$$= \lambda k.e^{\lambda(e^{-k}-1)-k}$$
, k being a real constant.

(Proved)

☆ Problem 7. If X_1 and X_2 be two independent Poisson random
(45) variables with common expected value θ, then show that the probability that $X_1 X_2$ is even is $\frac{1}{4}(3+2e^{-2\theta}-e^{-4\theta})$.

Soln. → $X_1 \sim P(\theta)$, $X_2 \sim P(\theta)$.

$$P(X_1X_2 = \text{even}) = 1 - P(X_1X_2 = \text{odd})$$
$$= 1 - \{P(X_1 = \text{odd}) \cdot P(X_2 = \text{odd})\} \quad \ldots (1) \quad [\because X_1, X_2 \text{ are indep.}]$$

Now,
$$P[X_1 = \text{odd}] = \theta \cdot e^{-\theta} + \frac{\theta^3 \cdot e^{-\theta}}{3!} + \frac{\theta^5 \cdot e^{-\theta}}{5!} + \cdots \infty$$
$$= e^{-\theta}\left[\theta + \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \cdots \infty\right]$$
$$= e^{-\theta}\left[\frac{e^{\theta} - e^{-\theta}}{2}\right]$$
$$= \frac{1 - e^{-2\theta}}{2}.$$

∴ (1) gives →

$$P(X_1X_2 = \text{even}) = 1 - \left(\frac{1 - e^{-2\theta}}{2}\right)\left(\frac{1 - e^{-2\theta}}{2}\right)$$
$$= 1 - \frac{1 - e^{-2\theta} - e^{-2\theta} + e^{-4\theta}}{4} \quad [\because P(X_1 = \text{odd}) = P(X_2 = \text{odd})]$$
$$= \frac{1}{4}\left(3 + 2e^{-2\theta} - e^{-4\theta}\right)$$

☆ Problem 8. If $X \sim P(\lambda)$ with mean $\mu = \lambda$ and variance $= \sigma^2 = \lambda$,
Show that :— $P[0 \le X \le 2\lambda] > 1 - \frac{1}{\lambda}$ (Using Chebyshev's inequality)

Soln. →
$X \sim P(\lambda)$
$\mu = \sigma^2 = \lambda$.

From Chebyshev's inequality, we know, —
$$P\left[\left|\frac{X-\mu}{\sigma}\right| \le t\right] > 1 - \frac{1}{t^2}, \text{ where } \mu, \sigma^2 < \infty.$$

So, $P[0 \le X \le 2\lambda] = P[|X - \mu| \le \lambda]$
$$= P[|X - \mu| \le \sqrt{\lambda} \cdot \sqrt{\lambda}]$$
$$= P\left[\left|\frac{X-\mu}{\sigma}\right| \le \sqrt{\lambda}\right] \quad [\because \sigma^2 = \lambda \text{ i.e. } \sigma = \sqrt{\lambda}]$$
$$> 1 - \frac{1}{(\sqrt{\lambda})^2} = 1 - \frac{1}{\lambda}.$$

Hence the proof is complete.

Problem 9. Let the p.m.f. f be positive on and only on the non-negative integers 0, 1, 2, . . . ; given that $f(x) = \frac{\lambda}{x} f(x-1)$ for $x = 1, 2, \ldots$

Determine 'f'.

(46)

soln. → $f(x) = \frac{\lambda}{x} f(x-1)$, $x = 1, 2, \ldots$

$f(1) = \lambda f(0)$

$f(2) = \frac{\lambda}{2} f(1) = \frac{\lambda^2}{2} f(0)$

$f(3) = \frac{\lambda}{3} f(2) = \frac{\lambda^3}{3!} f(0)$

$\vdots$

$f(x) = \frac{\lambda}{x} f(x) = \frac{\lambda^x}{x!} f(0)$ for $x = 1, 2, \ldots$

we know, $\sum_{x=0}^{\infty} f(x) = 1$ [$\because$ f is a p.m.f.]

$\Rightarrow \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \cdot f(0) = 1$

$\Rightarrow f(0) . e^{\lambda} = 1$

$\Rightarrow f(0) = e^{-\lambda}$.

$\therefore f(x) = \frac{\lambda^x . e^{-\lambda}}{x!}$ i.e. $X \sim P(\lambda)$.

Another way :— $\frac{f(x)}{f(x-1)} = \frac{\lambda}{x}$; $x = 1, 2, \ldots$

$= \frac{\frac{\lambda^x . e^{-\lambda}}{x!}}{\frac{\lambda^{x-1} . e^{-\lambda}}{(x-1)!}}$

$\therefore$ if $f(x) = \frac{\lambda^x . e^{-\lambda}}{x!}$, then $f(x-1) = \frac{\lambda^{x-1} . e^{-\lambda}}{(x-1)!}$, $x = 1, 2, \ldots$

If we can show $\sum_{x \in S} f(x) = 1$, then we can say, f(x) is the required p.m.f.

$\therefore \sum_{x=0}^{\infty} \frac{\lambda^x . e^{-\lambda}}{x!} = e^{-\lambda} . e^{\lambda} = 1$ (Checked).

So, $f(x) = \frac{\lambda^x . e^{-\lambda}}{x!}$; $X \sim P(\lambda)$.

Distribution function in terms of Incomplete Gamma function:

$X \sim P(\lambda)$ then for any non-negative integer k,

$$P[X \le k] = \frac{1}{\Gamma(k+1)} \int_{\lambda}^{\infty} e^{-u} u^k \, du \text{ . [incomplete Gamma function]}$$

__Proof:-__ $P[X \le k] = \sum_{x=0}^{k} \frac{e^{-\lambda} \cdot \lambda^x}{x!} = g(\lambda)$ .... (say)

$$g'(\lambda) = \frac{d}{d\lambda}(g(\lambda)) = \sum_{x=0}^{k} \frac{1}{x!}\left[e^{-\lambda} x \lambda^{x-1} - e^{-\lambda}\lambda^x\right]$$

$$= e^{-\lambda} \sum_{x=1}^{k} \frac{\lambda^{x-1}}{(x-1)!} - \sum_{x=0}^{k} e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x'=0}^{k-1} \frac{\lambda^{x'}}{x'!} - \sum_{x=0}^{k} e^{-\lambda} \cdot \frac{\lambda^x}{x!} \quad [(x-1) = x']$$

$$= -e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$\therefore \int_{\lambda}^{\infty} g'(u)\, du = \int_{\lambda}^{\infty} \frac{-e^{-u} u^k}{k!} du$$

$$\Rightarrow [g(u)]_{\lambda}^{\infty} = -\frac{1}{\Gamma(k+1)} \int_{\lambda}^{\infty} (e^{-u} \cdot u^k) du$$

$$\Rightarrow g(\lambda) = \frac{1}{\Gamma(k+1)} \int_{\lambda}^{\infty} e^{-u} \cdot u^k du .$$

$\therefore$ distribution function of X is a decreasing function of λ.

Now, $P[X > k] = 1 - P[X \le k]$

$$= 1 - \frac{1}{\Gamma(k+1)} \int_{\lambda}^{\infty} e^{-u} u^k du$$

$$= \frac{\int_{0}^{\infty} e^{-u} u^k du - \int_{\lambda}^{\infty} e^{-u} u^k du}{\Gamma(k+1)}$$

$$= \frac{1}{\Gamma(k+1)} \cdot \int_{0}^{\lambda} e^{-u} u^k du .$$

__Another way:__

$$P[X \le k] = \sum_{x=0}^{k} \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

$$= \frac{1}{\Gamma(k+1)} \sum_{x=0}^{k} \binom{k}{x} \Gamma(k-x+1) \cdot e^{-\lambda} \cdot \lambda^x$$

$$= \frac{1}{\Gamma(k+1)} \sum_{x=0}^{k} \binom{k}{x} e^{-\lambda} \cdot \lambda^x \int_{0}^{\infty} e^{-u} u^{k-x} du$$

$$= \frac{1}{\Gamma(k+1)} \int_0^\infty e^{-(\lambda+u)} \sum_{x=0}^{k} \binom{k}{x} \lambda^x u^{k-x} \, du$$

$$= \frac{1}{\Gamma(k+1)} \int_0^\infty (\lambda+u)^k \cdot e^{-(\lambda+u)} \, du$$

$$= \frac{1}{\Gamma(k+1)} \int_\lambda^\infty t^k e^{-t} dt \qquad \left[\begin{array}{l} \lambda+u=t \\ dt=du \end{array} \right]$$

$$= \frac{\int_0^\infty e^{-t} t^k dt - \int_0^\lambda e^{-t} t^k dt}{\lambda\,\Gamma(k+1)}$$

$$= 1 - \frac{\int_0^\lambda e^{-t} t^k dt}{\Gamma(k+1)}$$

☆ **Problem 10.** The r.v.'s x_1 & x_2 has Poisson distn. with parameter λ_1 & λ_2 respectively $\ni \lambda_1 < \lambda_2$. S.T.

$$P[X_1 \le k] > P[X_2 \le k].$$

Soln. → for any integer k, Distribution function of X, $\left\{ \frac{1}{\Gamma(k+1)} \int_\lambda^\infty e^{-u} u^k du \right\}$ or $\left\{ 1 - \frac{\int_0^\lambda e^{-u} u^k du}{\Gamma(k+1)} \right\}$ is a decreasing function λ.

So, if λ increases, the corresponding distribution function for any integer k decreases.

Here, $\lambda_1 < \lambda_2 \Rightarrow P[X_1 \le k] > P[X_2 \le k]$. (Proved)

☆ **Problem 11.** If $X \sim P(\lambda)$, then show that

$$P[X \geq n] < \frac{\lambda^n}{n!}.$$

Soln. → $P[X \geq n] = P[X > n-1] = \dfrac{\int_0^\lambda e^{-u} u^{n-1} du}{\Gamma(n)}$

$$< \frac{\int_0^\lambda u^{n-1} du}{\Gamma(n)} \quad [\because e^{-u} < 1]$$

$$= \frac{\lambda^n}{n!} \quad \text{(Proved)}$$

Reproductive property of Poisson Distribution :

(4a) A distn. is said to have the reproductive property if — $X_1, X_2, \ldots, X_n$ are independently distributed Poisson variates and $X_i \sim P(\lambda_i)$ for $i = 1(1)n$, then if $S_n = \sum_{i=1}^{n} X_i \sim P\left(\sum_{i=1}^{n} \lambda_i\right)$, then it is said to have the reproductive property.

Proof:-
$$P_{X_i}(t) = e^{-\lambda_i(1-t)}$$

Now, $P_{S_n}(t) = P_{X_1}(t) \cdot P_{X_2}(t) \cdots P_{X_n}(t)$ [$\because X_i$'s are indep.]

$$= e^{-\sum_{i=1}^{n} \lambda_i (1-t)}$$

$$\therefore \sum_{i=1}^{n} X_i \sim P\left(\sum_{i=1}^{n} \lambda_i\right)$$

$\therefore$ Clearly the reproductive property is additive.

Binomial distribution conditioning on Poisson variates :

Suppose X_1 & X_2 are independently distributed with Poisson variates $P(\lambda_1)$ & $P(\lambda_2)$, respectively. Then,

$$\{X / X_1 + X_2 = k\} \sim Bin\left(k, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right),$$

i.e. in case the poisson variates become i.i.d., we get $Bin(k, \frac{1}{2})$ distribution.

Distribution function of poisson distn. in terms of the distn. functn. of a continuous theoritical distribution :

If $X \sim P(\lambda)$.

Then, $$P[X \leq k] = 1 - \frac{\int_0^{\lambda} e^{-x} x^k dx}{\int_0^{\infty} e^{-x} x^k dx}$$

$$= 1 - \frac{\int_0^{\lambda} e^{-x} x^k dx}{\Gamma(k+1)}$$

Let, $$\frac{\int_0^{\lambda} e^{-x} x^k dx}{\Gamma(k+1)} = F(\lambda)$$

F is the distribution function of standard Gamma distribution with parameters $(k+1)$. $F(\lambda)$ is tabulated as $\Gamma\left(\frac{\lambda}{\sqrt{k+1}}, k\right)$ is incomplete gamma table as it is an incomplete gamma function.

☆ Problem 12. If $X \sim P(\lambda)$ and $Y/X = x \sim Bin(x, p)$, then S.T. $Y \sim P(\lambda p)$.

Soln. Here, $f(x) = \begin{cases} \dfrac{e^{-\lambda} . \lambda^x}{x!} & , x = 0, 1, 2, \ldots \\ 0 & \text{ow} \end{cases}$

& $f(y/x) = \begin{cases} \binom{x}{y} p^y q^{x-y} & , y = 0, 1, 2, \ldots, x \\ 0 & \text{ow} \end{cases}$

Hence, $f(y) = \sum_{x=y}^{\infty} f(y/x) . f(x)$

$$= \sum_{x=y}^{\infty} \binom{x}{y} p^y q^{x-y} . e^{-\lambda} . \frac{\lambda^x}{x!}$$

$$= \frac{e^{-\lambda} p^y}{y!} \sum_{x=y}^{\infty} \frac{\lambda^x . q^{x-y}}{(x-y)!}$$

$$= \frac{e^{-\lambda} . (\lambda p)^y}{y!} \sum_{x=y}^{\infty} \frac{\lambda^{x-y} . q^{x-y}}{(x-y)!}$$

$$= \frac{e^{-\lambda} . (\lambda p)^y}{y!} \sum_{u=0}^{\infty} \frac{(\lambda q)^u}{u!} \qquad [\text{taking } x-y=u]$$

$$= \frac{e^{-\lambda} . e^{\lambda q} . (\lambda p)^y}{y!}$$

$$= \frac{e^{-\lambda p} . (\lambda p)^y}{y!} \quad \text{for } y = 0, 1, 2, \ldots$$

Now, we can say, $\longrightarrow$ $y \sim P(\lambda p)$

Hence the proof is complete.

☆ Problem 13. Suppose the no. of eggs laid by an insect have a Poisson distribution with parameters λ and probability of eggs developing is p. Show that the no. of eggs surviving has also a Poisson distribution with parameter λp, assuming mutual independence of eggs.

(ci)

Soln. Let us define two random variables:

X denotes no. of eggs laid by an insect.

& Y denotes no. of insects born.

$y = 0, 1, \ldots,$

$$P[Y=y] = \sum_{x=0}^{\infty} P[X=x]\, P[Y=y \mid X=x]$$

$$= \sum_{x=y}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{x}}{x!} \times \binom{x}{y} p^{y} q^{x-y}$$

$$= \frac{e^{-\lambda} \lambda^{y} p^{y}}{y!} \sum_{x=y} \frac{(\lambda q)^{x-y}}{(x-y)!}$$

$$= \frac{e^{\lambda q} \cdot e^{-\lambda} (\lambda p)^{y}}{y!}$$

$$= \frac{e^{-\lambda p} (\lambda p)^{y}}{y!}$$

i.e. $Y \sim \text{Poi}(\lambda p)$

Scanned by CamScanner

☆ Problem 14. For an integer valued random variable X the p.m.f. $f(x)$ is such that →

$$f(x) = \frac{\alpha+\beta x}{x} \cdot f(x-1) \quad ; \text{where } x = 1, 2, 3, \ldots$$

where, $\beta \neq 1$.

Show that → $E(X) = \dfrac{\alpha+\beta}{1-\beta}$, $V(X) = \dfrac{\alpha+\beta}{(1-\beta)^2}$

$$MD_\mu(X) = \frac{2}{1-\beta}\left([\mu]+1\right) f\left([\mu]+1\right).$$

Soln : →

$$x f(x) = (\alpha+\beta x) f(x-1)$$

$$\Rightarrow \sum_{x=0}^{\infty} x f(x) = \alpha \sum_{x=1}^{\infty} f(x-1) + \beta \sum_{x=1}^{\infty} x f(x-1)$$

$$\Rightarrow E(X) = \alpha + \beta E(X+1) = \alpha + \beta E(X) + \beta$$

$$\Rightarrow E(X) = \frac{\alpha+\beta}{1-\beta}.$$

$$x^2 f(x) = \alpha x f(x-1) + \beta x^2 f(x-1)$$

$$\Rightarrow \sum_{x=0}^{\infty} x^2 f(x) = \alpha \sum_{x=1}^{\infty} x f(x-1) + \beta \sum_{x=1}^{\infty} x^2 f(x-1)$$

$$\Rightarrow E(X^2) = \alpha E(X+1) + \beta\left(E[X+1]^2\right)$$

$$\Rightarrow E(X^2) = \alpha E(X) + \alpha + \beta E(X^2) + 2\beta E(X) + \beta$$

$$\Rightarrow (1-\beta) E(X^2) = (\alpha+2\beta) E(X) + (\alpha+\beta)$$

$$\Rightarrow (1-\beta) E(X^2) = \frac{(\alpha+\beta)}{(1-\beta)} [\alpha+2\beta+1-\beta]$$

$$\Rightarrow E(X^2) = \frac{(\alpha+\beta)}{(1-\beta)^2} (\alpha+\beta+1)$$

$$\therefore Var(X) = E(X^2) - E^2(X)$$

$$= \frac{1}{(1-\beta)^2}\left[(\alpha+\beta)(\alpha+\beta+1) - (\alpha+\beta)^2\right]$$

$$= \frac{\alpha+\beta}{(1-\beta)^2}.$$

Given that → $$\frac{f(x)}{f(x-1)} = \frac{\alpha+\beta x}{x}$$

$$\Rightarrow \frac{f(x+1)}{f(x)} = \frac{\alpha+\beta(x+1)}{x+1}$$

$$\Rightarrow (x+1) f(x+1) = [\alpha+\beta(x+1)] f(x) \quad \cdots\cdots (1)$$

$$MD_\mu(X) = E|X-\mu|$$

$$= \sum_{x=0}^{\infty} \left| x - \frac{\alpha+\beta}{1-\beta} \right| f(x)$$

$$= 2 \sum_{x=0}^{n_0} \left(\frac{\alpha+\beta}{1-\beta} - x \right) f(x) \quad , \text{where } n_0 = [\mu]$$

$$= \frac{2}{1-\beta} \sum_{x=0}^{n_0} [\alpha+\beta-x+\beta x] f(x)$$

$$= \frac{2}{1-\beta}\left[\sum_{x=0}^{r_0}\{\alpha+\beta(x+1)\}f(x) - \sum_{x=1}^{r_0} x f(x)\right]$$

$$= \frac{2}{1-\beta}\left[\sum_{x=0}^{r_0}(x+1)f(x+1) - \sum_{x=1}^{r_0} x f(x)\right]$$

$$= \frac{2}{1-\beta}\left[\sum_{x'=1}^{r_0+1} x' f(x') - \sum_{x=1}^{r_0} x f(x)\right]$$

$$= \frac{2}{1-\beta}\left[(r_0+1)f(r_0+1)\right]$$

$$= \frac{2}{1-\beta}([\mu]+1)f([\mu]+1)$$

<u>Particular Cases</u>: $\rightarrow$ i) Let $X \sim Bin(n,p)$

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

$$\therefore \alpha = (n+1)\frac{p}{q} \text{ and } \beta = -\frac{p}{q}$$

$$\therefore E(X) = \frac{\alpha+\beta}{1-\beta} = \frac{\frac{np+p}{q} - \frac{p}{q}}{1+\frac{p}{q}} = np$$

$$\therefore V(X) = \frac{\alpha+\beta}{(1-\beta)^2} = \frac{\frac{np+p}{q} - \frac{p}{q}}{\left(1+\frac{p}{q}\right)^2} = npq.$$

ii) Let $X \sim P(\lambda)$

$$\therefore \frac{f(x)}{f(x-1)} = \frac{\lambda}{x}$$

$$\therefore \alpha = \lambda, \beta = 0$$

$$\therefore E(X) = \frac{\alpha+\beta}{1-\beta} = \lambda$$

$$\therefore V(X) = \frac{\alpha+\beta}{(1-\beta)^2} = \lambda$$

iii) Let $X \sim NB(r,p)$

$$\therefore \frac{f(x)}{f(x-1)} = \frac{r+x-1}{x}\cdot q$$

$$\therefore \alpha = (r-1)q, \quad \beta = q$$

$$\therefore E(X) = \frac{\alpha+\beta}{1-\beta} = \frac{(r-1)q+q}{1-q} = \frac{rq}{p}$$

$$\therefore V(X) = \frac{\alpha+\beta}{(1-\beta)^2} = \frac{(r-1)q+q}{(1-q)^2} = \frac{rq}{p^2}.$$

▨ GEOMETRIC DISTRIBUTION :—

<u>Definition</u>: A random variable X is said to follow a geometric distn. with parameter p if it has the p.m.f. of the following forms:

$$f(x) = \begin{cases} pq^{x} & , x = 0,1,2,\ldots \quad \text{(Model I)} \\ pq^{x-1} & , x = 1,2,3,\ldots \quad \text{(Model II)} \end{cases}$$

; where $0 < p < 1$ & $q = 1 - p$.

We denote, $X \sim Geo(p)$ or Geometric (p).

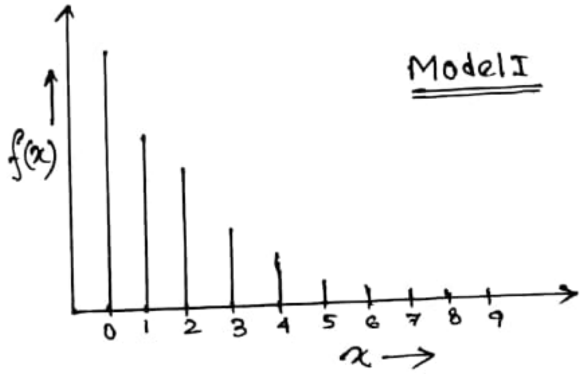


Figure showing column diagram, features of the distribution are quite evident from the column diagram where we see that the first mass point zero is the most probable value, i.e. the mode of the distn. and the distn. is positively skewed. $f(0)$ is maximum.

<u>Derivation of the pmf</u>:—

• <u>Probability model</u>:— Let us consider a sequence of Bernoulli trials with success probability p. Let the trials be repeated till the occurrance of the 1st success. Define:—

X : # failures preceeding the 1st success.

Y : # trials required to get the 1st success.

$\therefore P[X = x] = P[\, x$ failures have occured before the first success$]$

$= P[\underbrace{FF\ldots F}_{x}S] = pq^{x}$, $x = 0,1,2,\ldots\infty$; where $0<p<1$, $p+q=1$ ....(Model I)

$\therefore P[Y = x] = P[$ In the first $(x-1)$ trials failure has occured and in the last trial a success have occured $]$

$= P[\underbrace{FF\ldots F}_{x-1}S] = pq^{x-1}$, $x = 1,2,\ldots\infty$; where $0<p<1$, $q = 1-p$. ....(Model II)

i.e. $X \sim NB(1, p)$ or $X \sim Geo(p)$.

<u>Note</u>: An R.V. X that has a geometric distribution is also referred to as a <u>Discrete waiting time r.v.</u>, since it represents how long (in terms of failures/trials) one has to wait for the 1st success.

<u>Example</u>:→ Trials of keys required for opening a door.

Mean & Variance :–

We know, $\sum_{x=0}^{\infty} q^x = \frac{1}{1-q}$

Now, differentiating both sides w.r.t. q ;

$$\therefore \frac{d}{dq}\left[\sum_{x=0}^{\infty} q^x\right] = \frac{d}{dq}\left(\frac{1}{1-q}\right)$$

$$\Rightarrow \sum_{x=1}^{\infty} x q^{x-1} = \frac{1}{(1-q)^2} \quad \cdots\cdots (1)$$

$$\therefore E(X) = \sum_{x=0}^{\infty} x . pq^x$$

$$= \sum_{x=1}^{\infty} xpq^x \quad [\text{Using } (1)]$$

$$= \frac{pq}{(1-q)^2} = \frac{q}{p}.$$

Again, $\sum_{x=1}^{\infty} xq^{x-1} = \frac{1}{(1-q)^2}$

Differentiating both sides w.r.t. q ,

$$\Rightarrow \frac{d}{dq}\left[\sum_{x=1}^{\infty} x q^{x-1}\right] = \frac{d}{dq} \cdot \frac{1}{(1-q)^2}$$

$$\Rightarrow \sum_{x=2}^{\infty} x(x-1) q^{x-2} = \frac{2}{(1-q)^3} \quad \cdots\cdots (2)$$

$$\therefore E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) pq^x = \frac{2pq^2}{(1-q)^3} = \frac{2q^2}{p^2}.$$

$$\therefore V(X) = \frac{2q^2}{p^2} + \frac{q}{p} - \frac{q^2}{p^2} = \frac{q}{p^2}$$

Now, as $0 < p < 1$ and $p + q = 1$, so, —

$$\boxed{\text{Mean} < \text{Variance}}$$

PMF :–

$$P(t) = \sum_{x=0}^{\infty} t^x pq^x$$

$$= p \sum_{x=0}^{\infty} (qt)^x$$

$$= \frac{p}{(1-qt)}.$$

Factorial moment generating function :—

$$P(1+t) = p\left[1-q(1+t)\right]^{-1}$$
$$= p[p-qt]^{-1}$$
$$= \left[1-\frac{q}{p}t\right]^{-1}$$
$$= \sum_{x=0}^{\infty}\left(\frac{qt}{p}\right)^{x}$$

$$\mu'_{[x]} = E(X)_x$$
$$= \text{co-efficient of } \frac{t^x}{x!}$$
$$= x!\left(\frac{q}{p}\right)^{x}$$

raw moments

factorial moment

central moments

$$\therefore \mu'_1 = \mu'_{[1]} = \frac{q}{p}$$
$$\therefore \mu'_2 = \mu'_{[2]} + \mu'_1 = 2\left(\frac{q}{p}\right)^2 + \frac{q}{p}.$$

MGF :- $M(t) = \sum_{x=0}^{\infty} e^{tx}.pq^x = p\sum_{x=0}^{\infty}(qe^t)^x = \frac{p}{1-qe^t}.$

$$= p\left[1-q\left(1+t+\frac{t^2}{2!}+\cdots\infty\right)\right]^{-1}$$
$$= p\left[p-q\left(t+\frac{t^2}{2!}+\cdots\infty\right)\right]^{-1}$$
$$= \left[1-\frac{q}{p}\left(t+\frac{t^2}{2!}+\cdots\infty\right)\right]^{-1}$$
$$= 1+\frac{q}{p}\left(t+\frac{t^2}{2!}+\cdots\infty\right)$$
$$+\left(\frac{q}{p}\right)^2\left(t+\frac{t^2}{2!}+\cdots\infty\right)+\cdots$$

Mean = coefficient of $t = \frac{q}{p}$

$\mu_2' =$ coefficient of $\frac{t^2}{2!} = \frac{q}{p} + 2\left(\frac{q}{p}\right)^2$

$\therefore$ Variance $= 2\left(\frac{q}{p}\right)^2 + \frac{q}{p} - \frac{q^2}{p^2}$

$$= \frac{q}{p^2}.$$

Factorial moments :—

$$\mu_{[r]} = E[(X)_r] = \sum_{x=0}^{\infty}(x)_r\, pq^x$$
$$= pq^r(1-q)^{-r+1}.\, r!$$
$$= r!\left(\frac{q}{p}\right)^{r}.$$

Recursion Relation for central moments: —

(5)

$$\mu_r = \sum_{x=0}^{\infty}\left(x-\frac{q}{p}\right)^r . q^x p$$

$$= \sum_{x=0}^{\infty}\left(x-\frac{1-p}{p}\right)^r (1-p)^x . p$$

$$= \sum_{x=0}^{\infty}\left(x-\frac{1}{p}+1\right)^r (1-p)^x p$$

$$\therefore \frac{d\mu_r}{dp} = \sum_{x=0}^{\infty}\left\{ r\left(x-\frac{1}{p}+1\right)^{r-1} \cdot \frac{1}{p^2}(1-p)^x . p + \left(x-\frac{1}{p}+1\right)^r (1-p)^x \right.$$
$$\left. -\left(x-\frac{1}{p}+1\right)^r . x(1-p)^{x-1} . p \right\}$$

$$= \frac{r}{p^2}\mu_{r-1} + \frac{1}{p}\mu_r - \sum_{x=0}^{\infty}\left(x-\frac{1}{p}+1\right)^r\left(x-\frac{1}{p}+1+\frac{1}{p}-1\right)(1-p)^{x-1}.p$$

$$= \frac{r}{p^2}\mu_{r-1} + \frac{1}{p}\mu_r - \sum_{x=0}^{\infty}\left(x-\frac{1}{p}+1\right)^{r+1}(1-p)^{x-1}.p$$
$$- \sum_{x=0}^{\infty}\left(x-\frac{1}{p}+1\right)^r\left(\frac{1-p}{p}\right)(1-p)^{x-1} p$$

$$= \frac{r}{p^2}\mu_{r-1} + \cancel{\frac{1}{p}\mu_r} - \frac{1}{q}\mu_{r+1} - \cancel{\frac{\mu_r}{p}}$$

$$= \frac{r}{p^2}\mu_{r-1} - \frac{1}{q}\mu_{r+1}$$

$$\Rightarrow \boxed{\mu_{r+1} = q\left[\frac{r}{p^2}\mu_{r-1} - \frac{d\mu_r}{dp}\right]}$$

We know, $\mu_0 = 1$, $\mu_1 = 0$;

$r=1$, $\quad \mu_2 = q\left[\frac{1}{p^2}\times 1 - 0\right] = \frac{q}{p^2}$

$r=2$, $\quad \mu_3 = q\left[\frac{2}{p^2}\times 0 - \frac{d}{dp}\left(\frac{1-p}{p^2}\right)\right]$

$$= q\left[\frac{2}{p^3}(1-p) + \frac{1}{p^2}\right]$$

$$= q\left(\frac{2q+p}{p^3}\right)$$

$$= \frac{q(q+1)}{p^3}.$$

$r=3$, $\quad \mu_4 = q\left[\frac{3}{p^2}\mu_2 - \frac{d}{dp}\left\{\frac{(1-p)(2-p)}{p^3}\right\}\right]$

$$= q\left[\frac{3}{p^2}\cdot\frac{q}{p^2} - \frac{d}{dp}\left\{\frac{(1-p)(2-p)}{p^3}\right\}\right]$$

$$= q\left[\frac{3q}{p^4} - \left\{-3p^{-2}(1-p)(2-p) - (1-p)p^{-3}\right.\right.$$
$$\left.\left. -(2-p)p^{-3}\right\}\right]$$

$$\mu_4 = q\left[\frac{3q}{p^4} + 3p^{-2}q(2-p) + qp^{-3}(2-p)^{-3}.p^{-3}\right]$$

$$= \frac{3q^2}{p^4} + \frac{q}{p^2}$$

$$\therefore \mu_4 = \frac{q}{p^2}\left[3q/p^2 + 1\right]$$

Truncation of the mass point zero :⟶

Pmf of the truncated distribution $X/X>0$.

$$g(x) = \frac{P(x)}{P[X>0]} I_x\{1,2,\ldots\}$$

$$= \frac{pq^x}{1-P[X=0]} I_x\{1,2,\ldots\}$$

$$= \frac{pq^x}{1-p} I_x\{1,2,\ldots\}$$

$$= pq^{x-1} I_x\{1,2,\ldots\}$$

Let us consider a sequence of Bernoulli trials with probability of success p. Let the trials be repeated till the occurance of the first success.

Define :— X : No. of failures preceding the 1st success.
Y : No. of trials required to produce the 1st success.

$$\therefore P[X=x] = P[\underbrace{FF\ldots F}_{x}S] = pq^x I_x\{0,1,\ldots\}$$

$$P[Y=x] = P[\underbrace{FF\ldots F}_{x-1}S] = pq^{x-1} I_x\{1,2,\ldots\}$$

Now, $E(Y) = E(X+1)$

$$= E(X)+1$$

$$= \frac{q}{p}+1 = \frac{1}{p}$$

$$P[Y=y] = pq^{y-1} I_y\{1,2,\ldots\}$$

$$E(Y) = \sum_{x=1}^{\infty} x pq^{x-1}$$

$$= p[1+2q+3q^2+\ldots] = p(1-q)^{-2} = \frac{1}{p}.$$

$$\& \quad E[Y(Y+1)] = \sum_{x=1}^{\infty} x(x+1)pq^{x-1}$$

$$= \sum_{x=1}^{\infty} 2\binom{x+1}{2} pq^{x-1}$$

$$= \sum_{x'=0}^{\infty} 2\binom{x'+2}{2} pq^{x'} \qquad [\text{let, } x' = x-1]$$

$$= 2\sum_{x'=0}^{\infty} \binom{x'+3-1}{3-1} pq^{x'} = 2p(1-q)^{-3}$$

$$= \frac{2}{p^2}$$

$\therefore V(Y) = E[Y(Y+1)] - E(Y) - E^2(Y)$

$$= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{2-p-1}{p^2} = \frac{1-p}{p^2} = \frac{q}{p^2}$$

$\therefore Var(X) = Var(Y)$.

LOSS OF MEMORY PROPERTY :⟶

Geometric distribution lacks memory, i.e. if $X \sim Geo(p)$ then,–

$$P[X \geq i+j / X \geq i] = P[X \geq j] \quad, i,j = 0,1,\ldots$$

Here the converse is also true, i.e. the property is a characterisation of Geometric distribution.

And it is called the Loss of memory property of the geometric distribution. [WBSU'11]

Proof:–

If part :→ If $X \sim Geo(p)$ then, —

$$P[X \geq i+j / X \geq i] = P[X \geq j] \quad ; i,j = 0,1,2,\ldots$$

ANS:– $P[X \geq i] = \sum_{x=i}^{\infty} pq^x = p \cdot \frac{q^i}{1-q}$

$= q^i$,

$\therefore P[X \geq i+j] = q^{i+j}$

$\therefore P[X \geq i+j / X \geq i] = \frac{P[X \geq i+j]}{P[X \geq i]} = \frac{q^{i+j}}{q^i} = q^j$

$= P[X \geq j]$

$\therefore P[X \geq i+j] = P[X \geq i]\, P[X \geq j]$

$\therefore P[X \geq i+j / X \geq i] = P[X \geq j]$ (Proved)

Only if part :→ If $P[X \geq i+j / X \geq i] = P[X \geq j]$

then, → $X \sim Geo(p)$

ANS:–

$P_i = P[X \geq i]$

$\therefore P_{i+j} = P[X \geq i+j] = P[X \geq i]P[X \geq j]$

$= P_i \cdot P_j$

$P_{i+1} = P_1 \cdot P_i = P_1^2 P_{i-1} = \cdots = P_1^i \cdot P_{i-(i-1)}$

$= P_1^{i+1}$

$P_1 = P[X \geq 1] = 1 - P[X = 0]$

$\therefore P[X=0] = 1 - P_1$

$P_2 = P[X \geq 2] = 1 - P[X=0] - P[X=1]$

$\therefore P[X=1] = 1 - P_2 - (1-P_1)$

$= 1 - P_1^2 - 1 + P_1 = P_1(1-P_1)$

$P_3 = P[X \geq 3] = 1 - P[X=0] - P[X=1] - P[X=2]$

$\therefore P[X=2] = 1-(1-P_1) - P_1(1-P_1) - P_1^3$

$= \cancel{1} - \cancel{1} + \cancel{P_1} - \cancel{P_1} + P_1^2 - P_1^3$

$= P_1^2(1-P_1)$

Let it is true for $i = m-1$,

i.e. $P_{m+1} = P_1^{m+1}$

and $P[X = m-1] = P_1^{m-1}(1-P_1)$

$\therefore P[X \geq m+1] = 1 - \sum_{j=0}^{m} P[X=j]$

$P[X=m] = 1 - \sum_{j=0}^{m-1} P[X=j] - P_1^{m+1}$

$= P_1^m(1-P_1)$. (Proved)

$\therefore X \sim Geo(P_1)$.

EXAMPLE : $\longrightarrow$ As a crude example, if the arrival of a bus of a particular route at a specific bus stop in each minute (unit of time) is considered as a Bernoulli trial with a constant probability p, the probability that a person has to wait, at least for 5 minutes, remains independent of whether he has already waited for 3 minutes or, is just coming.

☆ Problem 1. Let $X \sim Geo(p)$, then show that — , $f(x) = pq^x$, $x = 0, 1, \ldots \infty$

$P[X > i+j] = P[X > i]P[X \geq j]$.

Proof:→ $P[X > i+j] = \sum_{x=i+j+1}^{\infty} pq^x = q^{i+j+1}$

Similarly, $P[X > i] = q^{i+1}$ and $P[X \geq j] = q^j$

$\therefore P[X > i]P[X \geq j] = q^{i+j+1}$

$\therefore P[X > i+j] = P[X > i]P[X \geq j] = P[X > i+j / X > i]$

or, $P[X > i+j / X > i] = P[X \geq j]$ . . . . ⊛

Note :→ in Problem 1, the Geometric distribution has no memory; ⊛ says that the information of no successes in i trials is forgotten in subsequent calculations.

The converse of the theorem is also true & discussed in the next page.

☆ Problem 2. Let X be a non-negative integer valued random variable satisfying the equation

(61)

$$P[X>i+j] = P[X>i]\,P[X\geq j] \quad \forall\, i,j = 1,2,\ldots\infty,$$

Then show that X must have a geometric distribution.

Proof:→ Let, $q_i = P[X>i]$, $i=0,1,\ldots\infty$

Then $q_{i+j} = P[X>i+j] = P[X>i]\,P[X\geq j]$

$= P[X>i]\,P[X>j-1]$

$= q_i\, q_{j-1}$

Taking $j=1$, we get ⟶

$$q_{i+1} = q_0 q_i = q_0^2 q_{i-1} = q_0^3 q_{i-2} = \cdots = q_0^{i+1} q_0 = q_0^{i+2}$$

$\therefore q_i = q_0^{i+1}$

$\therefore P[X=i] = P[X>i-1] - P[X>i]$

$= q_{i-1} - q_i$

$= q_0^i - q_0^{i+1}$

$= q_0^i(1-q_0)$

$\therefore X \sim Geo(p)$.

☆ Problem 3. Let X & Y be i.i.d. R.V.'s, and let

$$P[X=k] = p_k > 0, \quad k = 0,1,2,\ldots$$

if $P[X=t \mid X+Y=t] = P[X=t-1 \mid X+Y=t] = \frac{1}{1+t}$, $t\geq 0$

then X & Y are geometric RVs.

Soln → $P[X=t \mid X+Y=t] = P[X=t-1 \mid X+Y=t]$

$$\Rightarrow \frac{P[X=t,\, X+Y=t]}{P[X+Y=t]} = \frac{P[X=t-1,\, X+Y=t]}{P[X+Y=t]}$$

$\Rightarrow P[X=t, Y=0] = P[X=t-1, Y=1]$

$\Rightarrow P[X=t]\,P[Y=0] = P[X=t-1]\cdot P[Y=1]$

$\Rightarrow p_t\, p_0 = p_{t-1}\cdot p_1$

$\Rightarrow \frac{p_t}{p_{t-1}} = \frac{p_1}{p_0}$;

Hence, $p_t = \frac{p_t}{p_{t-1}} \times \frac{p_{t-1}}{p_{t-2}} \times \cdots\cdots \times \frac{p_1}{p_0} \times p_0$

$= \frac{p_1}{p_0} \times \frac{p_1}{p_0} \times \cdots\cdots \times \frac{p_1}{p_0} \times p_0$

$= \left(\frac{p_1}{p_0}\right)^t \cdot p_0$.

Since, $\sum_{t=0}^{\infty} p_t = 1$, we must have $p_1/p_0 < 1$.

Moreover, $p_0 \sum_{t=0}^{\infty} \left(\frac{p_1}{p_0}\right)^t = 1 \Rightarrow p_0 \cdot \frac{1}{1-\left(\frac{p_1}{p_0}\right)} = 1$

$\Rightarrow \frac{p_1}{p_0} = (1-p_0)$.

$\therefore p_t = (1-p_0)^t \cdot p_0$

$\therefore$ the proof is complete.

Another way:— We have,

$$P[X=t \mid X+Y=t] = \frac{p_t p_0}{\sum_{k=0}^{t} p_k p_{t-k}} = \frac{1}{t+1}.$$

and

$$P[X=t-1 \mid X+Y=t] = \frac{p_{t-1} \cdot p_1}{\sum_{k=0}^{t} p_k p_{t-k}} = \frac{1}{t+1}.$$

It follows that → $\frac{p_t}{p_{t-1}} = \frac{p_1}{p_0}$.

☆ Problem 4. Let X be a nonnegative integer-valued RV satisfying

$$P[X>m+1 \mid X>m] = P[X \geq 1]$$

for any non-negative integer m. then X must have a geometric distribution.

Proof:→ Let the PMF of X be written as

$$P[X=k] = p_k, \quad k=0,1,2,\ldots$$

then $P[X \geq n] = \sum_{k=n}^{\infty} p_k$

and $P[X>m] = \sum_{k=m+1}^{\infty} p_k = q_m$; say,

$$P[X>m+1 \mid X>m] = \frac{P[X>m+1]}{P[X>m]} = \frac{q_{m+1}}{q_m}.$$

Thus, $q_{m+1} = q_m q_0$.

where $q_0 = P[X>0] = p_1 + p_2 + \cdots = 1 - p_0$. It follows that

$$q_k = (1-p_0)^{k+1},$$

and hence $p_k = q_{k-1} - q_k = (1-p_0)^k p_0$, as asserted.

∴ The proof is complete.

Interpretation of Loss of memory property:— Probability that more than m+n trials will be required before the first success given that there have been already more than m failures is equal to the unconditional probability that at least n trials are needed before the 1st success. Therefore, the information of no. of successes in (m+1) trials is forgotten in subsequent calculation.

☆ **Problem 5.** Express negative binomial distribution as a sum of a number of geometric distribution.

Soln: → X can be written as, $X = X_1 + X_2 + \cdots + X_r$, where

X_i = No. of failures preceeding the ith success after having $(i-1)$ success.

X_i's are i.i.d. geometric random variable having p.m.f.

$$P[X_i = x] = pq^x \, I_x\{0, 1, \ldots\}$$

PGF of X is given by,

$$P_X(t) = [P_{X_i}(t)]^r \quad ; \quad P_{X_1}(t) = \sum_{x=0}^{\infty} t^x . pq^x = (1-qt)^{-1}.p$$

$$= p^r (1-qt)^{-r}$$

$$= \sum_{x=0}^{\infty} \frac{(-r)x}{x!} p^r . (-qt)^x$$

$$P[X = x] = \text{coefficient of } t^x$$

$$= \frac{(-r)x}{x!} (-q)^x p^r$$

$$= \binom{x+r-1}{x} p^r q^x \quad , ///$$

$$E(X) = E\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} E(X_i) = \sum_{i=1}^{r} \frac{q}{p} = \frac{rq}{p}.$$

$$V(X) = V\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} V(X_i) + 0$$ [covariance term vanishes due to independence]

$$= \sum_{i=1}^{r} \frac{q}{p^2} = \frac{rq}{p^2}.$$

Another method: →

Similarly, we can write,

$N = N_1 + N_2 + \cdots + N_r$ = No. of trials required to produce r success.

N_j = No. of trials required to produce jth success after having $j-1$ success.

Clearly, N_j's are i.i.d. random variable with p.m.f.

$$P[N_j = x] = P[\underbrace{FF\cdots F}_{x-1}S] = pq^{x-1}, \; x = 1, 2, \ldots \quad \forall j$$

P.G.F. of N_j is,

$$E(t^{N_j}) = \sum_{x=1}^{\infty} t^x pq^{x-1} = pt \sum_{x=1}^{\infty} (qt)^{x-1}$$

$$= pt(1-qt)^{-1} \quad \left[\because |t| < \frac{1}{q}\right]$$

$$E(t^N) = \text{P.G.F of } N$$
$$= [E(t^{N_1})]^{r}$$
$$= (pt)^r (1-qt)^{-r} \quad [\because \text{ they are i.i.d.}]$$
$$= (pt)^r \sum_{i=0}^{\infty} \frac{(-r)_i}{i!} (-qt)^i$$
$$= \sum_{i=0}^{\infty} \binom{r-1+i}{i} p^r q^i t^{r+i}$$
$$= \sum_{i'=r}^{\infty} \binom{i'-1}{r-1} p^r q^{i'-r} \cdot t^{i'}$$

Coefficient of $t^m = P[N=m] = \binom{m-1}{r-1} p^r q^{m-r} I_m\{r, r+1, \ldots\}$

.///

$$E(N) = \sum_{i=1}^{r} E(N_i)$$
$$= \sum_{i=1}^{r} \frac{1}{p}$$
$$= \frac{r}{p}$$

$$V(N) = V\left(\sum_{i=1}^{r} N_i\right)$$
$$= \sum_{i=1}^{r} V(N_i) + 0$$
$$= \frac{rq}{p^2}$$

[covariance term vanishes due to independence]

Polya's Urn Scheme / Polya's Distribution :→

Suppose an urn contains 'a' white and 'b' black balls. Now the balls are drawn at random one by one with replacement, moreover everytime 'c' balls of the same colour and 'd' balls of opposite colour are added; this is known as Polya's urn scheme.

While adopting clinical trials to ethically justify the design adoptive allocations are usually made. Such as, play-the-winner-rule and randomise-play the winner rule. These designs are directly designed from Polya's urn scheme.

Polya's Distribution : According to the urn scheme n random drawings are made,

Let, X = Number of white balls drawn

For simplicity we discuss the case $d=0$

$$\therefore P[X=x] = \binom{n}{x} P[\omega_1 \omega_2 \cdots \omega_x \omega_{x+1}^c \cdots \omega_n^c]$$

$$= \binom{n}{x} \frac{a}{a+b} \cdot \frac{a+c}{a+b+c} \cdots \frac{a+(x-1)c}{a+b+(x-1)c} \cdot \frac{b}{a+b+xc} \cdots \frac{b+(n-x-1)c}{a+b+(n-1)c}$$

[$\because$ there are $\binom{n}{x}$ mutually exclusive and equally likely cases]

Put, $a=-\alpha c$
$b=-\beta c$

$$= \binom{n}{x} \frac{\alpha(\alpha-1)\cdots(\alpha-x+1)\beta(\beta-1)\cdots(\beta-n+x+1)}{(\alpha+\beta)(\alpha+\beta-1)\cdots(\alpha+\beta-n+1)}$$

$$= \binom{n}{x} \cdot \frac{(\alpha)_x (\beta)_{n-x}}{(\alpha+\beta)_n}$$

$$= \frac{\frac{(\alpha)_x}{x!} \cdot \frac{(\beta)_{n-x}}{(n-x)!}}{\frac{(\alpha+\beta)_n}{n!}}$$

$$= \frac{\binom{\alpha}{x}\binom{\beta}{n-x}}{\binom{\alpha+\beta}{n}}.$$

Thus the p.m.f. of the distribution can be written in terms of the hypergeometric p.m.f.

$$E(X) = \frac{n\alpha}{\alpha+\beta}$$

$$V(X) = \frac{(\alpha+\beta-n)}{(\alpha+\beta-1)} \cdot n \cdot \frac{\alpha\beta}{(\alpha+\beta)^2}$$

[As per Hypergeometric distributions mean & variance]

in terms of a and b;

$$E(X) = \frac{an}{a+b}.$$

$$V(X) = \frac{a+b+nc}{a+b+c} \times \frac{nab}{(a+b)^2}$$

DEGENERATE DISTRIBUTION :—

An r.v. is said to be degenerate at 'c' if $P[X=c]=1$, i.e. $P[X=x] = \begin{cases} 1 & \text{if } x=c \\ 0 & \text{ow} \end{cases}$

Cumulative Distribution function :—

$$F_X(x) = P[X \le x] = \begin{cases} 0 & \text{if } x < c, \\ 1 & \text{if } x \geq c \end{cases}$$

Moments :—

$$\mu_r' = c^r \cdot P[X=c] + 0 = c^r.$$

$$\therefore Var(X) = \mu_2' - \mu_1'^2 = c^2 - c^2 = 0$$

Var(X) = 0 is a characterisation of a degenerate R.V.

$Var(X) = 0$, $\mu = E(X)$.

$\Leftrightarrow E(X-\mu)^2 = 0$

$\Leftrightarrow P[X-\mu=0] = 1$

$\Leftrightarrow P[X=\mu] = 1.$

HYPERGEOMETRIC & WAITING TIME HYPERGEOMETRIC (INVERSE HYPERGEOMETRIC) DISTRIBUTION :→

Probability Model:— Consider a box containing 'N' objects of which 'M' are of one kind (say, type A) and N-M are of other kind (say, type A^c).

Suppose n objects are drawn at random from the box.

X = No. of type A objects drawn in the sample,

Then the p.m.f. of X is

$$P[X=x] = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & ; \text{where } x \in [\max(0, n+M-N), \min(M,n)] \\ 0 & \text{ow} \end{cases}$$

Note,— $0 \le x \le M$

$0 \le n-x \le N-M$

$\Rightarrow n+M-N \le x \le n$

$\therefore x \in \max(0, n+M-N), \ldots\ldots, \min(M,n)$.

If the R.V. X follows a hypergeometric distribution with parameters n, N, M;

we write → $X \sim H(N,n,M)$.

Let, $p = \frac{M}{N}$ = Proportion of type A objects in the box.

$q = \frac{N-M}{N} = 1-p$.

$$\therefore f(x) = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} \quad \text{if } x = 0, 1, \ldots, \min(n, Np), \; p, q > 0, \; p+q=1.$$

$= 0$ ow

We write, $X \sim H(n, N, p)$

In practice n is so chosen that, $n \le \min(Np, Nq)$.

Then the mass points of X are $0, 1, 2, \ldots, n$.

Result:- Suppose an urn contains 'a' white and 'b' black balls. Consider the following scheme (unique drawing method)

(i) n balls are drawn at random one by one with replacement.

(ii) balls are drawn at random one by one with replacement till n white balls being produced.

(iii) n balls are drawn at random one by one without replacement.

(iv) balls are drawn at random one by one without replacement till n white balls being produced.

Proof:- (i) & (iii) X : No. of white balls drawn.

(ii) & (iv) Y : No. of black balls preceeding the nth white ball.

N : No. of drawings necessary to produce n white balls.

(i) $P[X=x] = \binom{n}{x} P[\omega_1 \omega_2 \cdots \omega_x \omega_{x+1}^c \cdots \omega_n^c]$ [$\because$ there are $\binom{n}{x}$ mutually exclusive and equally probable cases]

$= \binom{n}{x} P(\omega_1)P(\omega_2)\cdots P(\omega_x)P(\omega_{x+1}^c) \cdots P(\omega_n^c)$ [$\because$ they are independent]

$= \binom{n}{x}\left(\frac{a}{a+b}\right)^x \left(\frac{b}{a+b}\right)^{n-x} I_x\{0,1,\ldots,n\}$

$\therefore X \sim Bin\left(n, \frac{a}{a+b}\right)$.

(ii) $P[Y=x] = \binom{n-1+x}{x} P[\omega_1\omega_2\cdots\omega_{n-1}\cdot\omega_n^c\cdots\omega_{n+x-1}^c\,\omega_{n+x}]$

$= \binom{n-1+x}{x}\left(\frac{a}{a+b}\right)^n\left(\frac{b}{a+b}\right)^x I_x\{0,1,\ldots\infty\}$ [$\because$ there are $\binom{n-1+x}{x}$ m.e. and equally probable cases]

$\therefore Y \sim NB\left(n, \frac{a}{a+b}\right)$

$P[N=x] = \binom{x-1}{n-1} P[\omega_1,\omega_2\cdots w_{n-1} w_n^c \cdots w_{x-1}^c\, w_x]$

$= \binom{x-1}{n-1}\left(\frac{a}{a+b}\right)^n\left(\frac{b}{a+b}\right)^{x-n} I_x\{1,2,\ldots\infty\}$ [$\because$ there are $\binom{x-1}{n-1}$ m.e. & e.p. cases]

$\therefore Y \sim NB\left(n, \frac{a}{a+b}\right)$

(iii) $P[X=x] = \binom{n}{x} P(W_1)\, P(W_2/W_1)\, P(W_3/W_1W_2) \cdots$

$P(W_x/W_1\cdots W_{x-1})\, P(W_{x+1}^c / W_1\cdots W_x) \cdots$

$\cdots P(W_n^c / W_1 \cdots W_x W_{x+1}^c \cdots W_{n-1}^c)$

$$= \binom{n}{x} \frac{a(a-1)\cdots(a-x+1)\, b(b-1)\cdots(b-\overline{n-x}+1)}{(a+b)(a+b+1)\cdots\cdots(a+b-n+1)}$$

$$= \binom{n}{x} \frac{(a)_x\, (b)_{n-x}}{(a+b)_n}$$

$$= \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}\, I_x\{\max(0, n-b), \ldots, \min(n,a)\}$$

[This is Hypergeometric Distribution]

(iv) $P[Y=x] = \binom{n-1+x}{x} P(W_1)\, P(W_2/W_1) \cdots P(W_{n-1}/W_1 \cdots W_{n-2})$

$P(W_n^c / W_1 \cdots W_n) \cdots$

$P(W_{n+x-1}^c / W_1 \cdots W_{n-1} W_n^c \cdots W_{n+x-2}^c)$

$P(W_{n+x} / W_1 \cdots W_{n-1} W_n^c \cdots W_{n+x-1}^c)$

$$= \binom{n-1+x}{x} \frac{a(a-1)\cdots(a-n+2)\, b(b-1)\cdots(b-x+1)(a-n+1)}{(a+b)(a+b-1)(a+b-2)\cdots(a+b-\overline{n+x}+1)}$$

$$= \binom{n-1+x}{x} \frac{(a)_n\, (b)_x}{(a+b)_{n+x}}\, I_x\{0, \ldots, b\}$$

[This is waiting time hypergeometric distribution]

OR. Let us define the events A & B.

A = $(n-1)$ whites balls come out in $(n-1+x)$ drawings.

B = one white ball come out in the $(n+x)$th drawing.

$P[Y=x] = P[A\cap B] = P(A)\, P(B/A)$

$$= \binom{n-1+x}{n-1} \frac{(a)_{n-1}\, (b)_x}{(a+b)_{n+x-1}} \times \frac{a-n+1}{a+b-n-x+1}$$

$$= \binom{n+k-1}{n-1} \frac{(a)_n\, (b)_x}{(a+b)_{n+x}}$$ [Negative hypergeometric distribution]

Now,

$P[N=x] = \binom{x-1}{n-1} P(W_1)\, P(W_2/W_2) \cdots P(W_{n-1}/W_1 \cdots W_{n-2})$

$P(W_n^c / W_1 \cdots W_{n-1}) \cdots P(W_{x-1}^c / W_1 \cdots W_{n-1} W_n^c$

$\cdots W_{x-2}^c)\, P(W_x / W_1 \cdots W_{n-1} W_n^c \cdots W_{x-1}^c)$

$$= \binom{x-1}{n-1} \frac{a(a-1)\cdots(a-n+2)\, b(b-1)\cdots(b-x+n+1)(a-n+1)}{(a+b)(a+b-1)\cdots\cdots(a+b-x+1)}$$

$$= \binom{x-1}{n-1} \frac{(a)_n\, (b)_{x-n}}{(a+b)_x}\, I_x\{n, n+1, \ldots, n+b\}$$

Moments of Hypergeometric & Waiting time Hypergeometric distribution :

If $X \sim$ Hypergeometric $(a+b, a, n)$

① $\therefore P[X=x] = \dfrac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$, $x = \{\max(0, n-b), \ldots, \min(n, a)\}$

Now, the rth order factorial moment will be defined as —

$$\mu_{[k]} = E(X)_k = \sum_{x=0}^{n} (x)_k \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$$

$$= \sum_{x=k}^{n} (x)_k \cdot \frac{a!}{x!(a-x)!} \cdot \frac{\binom{b}{n-x}}{\binom{a+b}{n}}$$

$$= (a)_k \sum_{x=k}^{n} \frac{(a-k)!}{(x-k)!(a-\cancel{k}-x-k)!} \times \frac{\binom{b}{n-x}}{\binom{a+b}{n}}$$

$$= (a)_k \sum_{x=k}^{n} \frac{\binom{a-k}{x-k}\binom{b}{n-x}}{\binom{a+b}{n}}$$

$$= (a)_k \sum_{x'=0}^{n-k} \frac{\binom{a-k}{x'}\binom{b}{n-x'-k}}{\frac{(a+b)!}{n!(a+b-n)!}}$$

$$= (a)_k \sum_{x'=0}^{n-k} \frac{\binom{a-k}{x'}\binom{b}{n-x'-k}}{\frac{(a+b)_k (a+b-k)!}{(n)_k (n-k)!(a+b-k-n+k)!}}$$

$$= \frac{(a)_k (n)_k}{(a+b)_k} \sum_{x'=0}^{n-k} \frac{\binom{a-k}{x'}\binom{b}{n-x'-k}}{\binom{a+b-k}{n-k}}$$

$$= \frac{(a)_k (n)_k}{(a+b)_k}$$

$\therefore E(X) = \dfrac{na}{a+b}$, $E[X(X-1)] = \dfrac{n(n-1)a(a-1)}{(a+b)(a+b-1)}$

$\therefore V(X) = E(X^2) - E^2(X)$

$$= E(X(X-1)) + E(X) - E^2(X)$$

$$= \frac{n(n-1)a(a-1)}{(a+b)(a+b-1)} + \frac{na}{(a+b)} - \frac{n^2a^2}{(a+b)^2}$$

$$= \frac{a+b-n}{a+b-1} \cdot \frac{nab}{(a+b)^2}$$

(Waiting time hypergeometric Distribution).

② $$P[Y=x] = \binom{-n}{x}(-1)^x \cdot \frac{(a)_n\,(b)_x}{(a+b)_{n+x}}, \quad x=0,1,2,\dots,b$$

$$E(Y)_k = \sum_{x=k}^{b} (x)_k \binom{-n}{x}(-1)^x \frac{(a)_n\,(b)_x}{(a+b)_{n+x}}$$

$$= (-n)_k(-1)^k \sum_{x=k}^{b} \binom{-n-k}{x-k}(-1)^{x-k}\frac{(a)_n\,(b)_x}{(a+b)_{n+x}}$$

$$= \frac{(-n)_k(-1)^k(b)_k}{(a+k)_k}\sum_{x=k}^{b} \binom{-n-k}{x-k}(-1)^{x-k}\cdot\frac{(a+k)_{n+k}\,(b-k)_{x-k}}{(a+k+b-k)_{n+k+x-k}}$$

$$= \frac{(-n)_k(-1)^k(b)_k}{(a+k)_k}\sum_{x'=0}^{b-k}\binom{-\overline{n+k}}{x'}(-1)^{x'}\cdot\frac{(a+k)_{n+k}\,(b-k)_{x'}}{(a+b)_{n+k+x'}}$$

$$= \frac{(-n)_k(-1)^k(b)_k}{(a+b)_k}$$

$\therefore E(Y) = \dfrac{nb}{a+1}$; $\quad E[Y(Y-1)] = \dfrac{n(n+1)\,b(b-1)}{(a+1)(a+2)}$

$$\therefore V(Y) = E(Y^2) - E^2(Y)$$

$$= \frac{n(n+1)\,b(b-1)}{(a+1)(a+2)} + \frac{nb}{(a+1)} - \left(\frac{nb}{a+1}\right)^2$$

③ $$P[N=x] = \binom{x-1}{n-1}\frac{(a)_n\,(b)_{x-n}}{(a+b)_x} \quad \text{for } x=n,\dots,b+n$$

$$E[N(N+1)(N+2)\cdots(N+k-1)]$$

$$= \sum_{x=n}^{b+n} x(x+1)\cdots(x+k-1)\binom{x-1}{n-1}\frac{(a)_n\,(b)_{x-n}}{(a+b)_x}$$

$$= \cfrac{n(n+1)\cdots(n+k-1)}{\cfrac{(a+k)_k}{(a+b+k)_k}}\sum_{x=n}^{b+n}\binom{x+k-1}{n+k-1}\frac{(a+k)_{n+k}\,(b)_{x-n}}{(a+b+k)_{x+k}}$$

$$= \frac{(a+b+k)_k\; n(n+1)\cdots(n+k-1)}{(a+k)_k}$$

$$\therefore E(N) = \frac{(a+b+1)n}{(a+1)}$$

$$\therefore V(N) = E(N(N+1)) - E(N) - E^2(N) = \frac{(a+b+2)(a+b+1)n(n+1)}{(a+1)(a+2)} - \frac{(a+b+1)n}{(a+1)} - \left[\frac{(a+b+1)n}{(a+1)}\right]^2$$

Again, we know,

$N = Y + n$

$$\therefore E(N) = E(Y) + n = \frac{nb}{a+1} + n = n\left[\frac{a+b+1}{a+1}\right]$$

and, $V(N) = V(Y)$.

Alternative Approach for calculating factorial moments :→

Factorial moments : → $f(x) = \dfrac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$.

$$\mu_{[k]} = E(X)_k = \sum_{x=k}^{n} (x)_k \cdot \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=k}^{n} (x)_k \cdot \frac{(Np)!}{x!\,(Np-x)!} \cdot \binom{Nq}{n-x}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=k}^{n} \frac{(Np)!}{(x-k)!\,(Np-x)!} \binom{Nq}{n-x}$$

$$= \frac{1}{\binom{N}{n}} (Np)_k \sum_{x=k}^{n} \binom{Np-k}{x-k}\binom{Nq}{n-x}$$

$$= \frac{1}{\binom{N}{n}} (Np)_k \sum_{x'=0}^{n-k} \binom{Np-k}{x'}\binom{Nq}{n-k-x'}$$

Note, $(1+t)^{Np-k} \cdot (1+t)^{Nq} = (1+t)^{N-k}$

$$\Rightarrow \left\{ \cdots + \binom{Np-k}{x} t^{x} + \cdots \right\} \left\{ \cdots + \binom{Nq}{n-k-x} t^{n-k-x} + \cdots \right\}$$

$$= \left\{ \cdots + \binom{N-k}{n-k} t^{n-k} + \cdots \right\}$$

Collecting coefficient of t^{n-k} from both side we get, —

$$\sum_{x=0}^{n-k} \binom{Np-k}{x}\binom{Nq}{n-k-x} = \binom{N-k}{n-k}$$

$$\therefore \mu_{[k]} = \frac{(Np)_k}{\binom{N}{n}} \binom{N-k}{n-k}$$

$$\therefore \mu_{[1]} = E(X) = \frac{Np\binom{N-1}{n-1}}{\binom{N}{n}} = np$$

$$\therefore \mu_{[2]} = E[X(X-1)]$$

$$= \frac{Np(Np-1)}{\binom{N}{n}} \cdot \binom{N-2}{n-2} = \frac{(Np-1)(n-1)np}{(N-1)}$$

$$\therefore \mu_2 = var(X) = \frac{p(Np-1)n(n-1)}{N-1} + np - (np)^2 = \frac{npq(N-n)}{N-1}.$$

$$\therefore \mu_3 = npq(q-p)\frac{(N-n)(N-2n)}{(N-1)(N-2)}$$

$$\therefore \mu_4 = \frac{npq(N-n)}{(N-1)(N-2)(N-3)} \cdot \left[N(N+1) - 6n(N-n) + pq\left\{N^2(n-2) - Nn^2 + 6n(N-n)\right\}\right]$$

Mode of the Hypergeometric Distribution: —

$$\frac{f(x)}{f(x-1)} = \frac{\binom{Np}{x}\binom{Nq}{n-x}/\binom{N}{n}}{\binom{Np}{x-1}\binom{Nq}{n-x+1}/\binom{N}{n}} = \frac{n-x+1}{x}\cdot\frac{(Np-x+1)}{(Nq-n+x)}$$

$\therefore f(x) \gtreqless f(x-1)$ according as $\dfrac{(n-x+1)(Np-x+1)}{x(Nq-n+x)} \gtreqless 1$

or, $Npn - Npx + Np - nx + x^2 - x + n - x + 1 \gtreqless xNq - nx + x^2$

or, $Npn - Nx - 2x + Np + n + 1 \gtreqless 0$

or, $Npn + Np + n + 1 \gtreqless (N+2)x$

or, $(Np+1)(n+1) \gtreqless (N+2)x$

$\Rightarrow x \lesseqgtr \dfrac{(Np+1)(n+1)}{(N+2)}$.

$\therefore x$ has mode at $\hat{x} = \left[\dfrac{(Np+1)\quad(n+1)}{(N+2)}\right]$ if x is not an integer.

if x be an integer then two modes are at

$\dfrac{(Np+1)(n+1)}{(N+2)} - 1$ and $\dfrac{(Np+1)(n+1)}{(N+2)}$.

Binomial Approximation to Hypergeometric Distribution : —

If $X \sim H.(N, n, p)$ and $N >> n$ then $\frac{n}{N}$ can be ignored, i.e. $\frac{n}{N} \to 0$ as $N \to \infty$ and n = fixed. Then —
the Hypergeometric distribution will converge to a binomial distribution with parameters n & p.

Proof :→ For some fixed x

$$P[X=x] = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} = \frac{\frac{(NP)_x}{N^x}\cdot\frac{(Nq)_{n-x}}{N^{n-x}}\binom{n}{x}}{\frac{(N)_n}{N^n}}$$

$$= \frac{\binom{n}{x}\left(\frac{NP}{N}\right)\left(\frac{Np-1}{N}\right)\cdots\left(\frac{Np-\overline{x-1}}{N}\right)\left(\frac{Nq}{N}\right)\left(\frac{Nq-1}{N}\right)\cdots\left(\frac{Nq-\overline{n-x-1}}{N}\right)}{\left(\frac{N}{N}\right)\left(\frac{N-1}{N}\right)\left(\frac{N-2}{N}\right)\cdots\left(\frac{N-\overline{n-1}}{N}\right)}$$

$$= \binom{n}{x}\frac{p\left(p-\frac{1}{N}\right)\left(p-\frac{2}{N}\right)\cdots\left(p-\frac{x-1}{N}\right)q\left(q-\frac{1}{N}\right)\cdots\left(q-\frac{n-x-1}{N}\right)}{1\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)}$$

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

as $N \to \infty$

$p - \frac{i}{N} \to p \quad \forall\ i = 1(1)x-1$

$q - \frac{j}{N} \to q \quad \forall\ j = 1(1)n-x-1$

Interpretation : — If $N \to \infty$, the popl^n. become an infinite popl^n. From an infinite popl^n., drawing samples WR and WOR are practically same, since we may have a sampling units repeatedly under WOR. Then $P[X=x]$ under WOR tends to $P[X=x]$ under WR as $N \to \infty$.

Mean Deviation about Mean :→

N — items $= a+b$

Np (Type A) $\quad$ Nq (Type A^c)

$$p = \frac{a}{N} \qquad p+q = \frac{a+b}{N} = 1$$

$$q = \frac{b}{N} \qquad 0 \le p \le 1.$$

n elements are drawn without replacement

$$P[X=x] = \binom{Np}{x}\binom{Nq}{n-x} \Big/ \binom{N}{n}$$

Mean deviation about mean is defined as ⟶

$$E|X-np|$$

$$= \frac{2}{\binom{N}{n}} \sum_{x=n_0}^{n} (x-np)\binom{Np}{x}\binom{Nq}{n-x} \quad \text{, where } n_0 = [np+1]$$

$$= \frac{2}{\binom{N}{n}} \sum_{x=n_0}^{n} [x(p+q)-np]\binom{Np}{x}\binom{Nq}{n-x}$$

$$= \frac{2}{\binom{N}{n}} \sum_{x=n_0}^{n} [xq-(n-x)p]\binom{Np}{x}\binom{Nq}{n-x}$$

$$= \frac{2}{\binom{N}{n}} \left[q \sum_{x=n_0}^{n} \cancel{x} \cdot \frac{Np}{\cancel{x}} \binom{Np-1}{x-1} \left\{ \binom{Nq-1}{n-x} + \binom{Nq-1}{n-x-1} \right\} - p \sum_{x=n_0}^{n} \cancel{(n-x)} \frac{Nq}{\cancel{(n-x)}} \binom{Nq-1}{n-x-1} \left\{ \binom{Np-1}{x} + \binom{Np-1}{x-1} \right\} \right]$$

$$= \frac{2Npq}{\binom{N}{n}} \left[\sum_{x=n_0}^{n} \binom{Np-1}{x-1}\binom{Nq-1}{n-x} + \sum_{x=n_0}^{n-1} \cancel{\binom{Np-1}{x-1}\binom{Nq-1}{n-x-1}} - \sum_{x=n_0}^{n-1} \binom{Nq-1}{n-x-1}\binom{Np-1}{x} - \sum_{x=n_0}^{n-1} \cancel{\binom{Nq-1}{n-x-1}\binom{Np-1}{x-1}} \right]$$

$$= \frac{2Npq}{\binom{N}{n}} \left[\sum_{x'=n_0-1}^{n-1} \binom{Np-1}{x'}\binom{Nq-1}{n-x'-1} - \sum_{x=n_0}^{n-1} \binom{Np-1}{x}\binom{Nq-1}{n-x-1} \right]$$

$$= \frac{2Npq}{\binom{N}{n}} \binom{Np-1}{n_0-1}\binom{Nq-1}{n-n_0}$$

$$= \frac{2\cancel{Npq}}{\binom{N}{n}} \times \frac{n_0}{\cancel{Np}} \binom{Np}{n_0} \times \frac{Nq-n+n_0}{\cancel{Nq}} \binom{Nq}{n-n_0}$$

$$= \frac{2n_0(Nq-n+n_0)}{N} \cdot P[X=n_0]$$

Alternative way :— Let $X \sim H(N, n, p)$.

$\therefore E(X) = np = \mu$.

$$MD_\mu(X) = \sum_{x=0}^{n} |x-\mu| f(x)$$

$$= 2\sum_{x=0}^{n_0} (\mu - x) f(x) \qquad \text{where } n_0 = [\mu]$$

$$= 2\sum_{x=0}^{n_0} [np - x(p+q)] f(x)$$

$$= 2\sum_{x=0}^{n_0} [p(n-x) - qx] f(x)$$

$$= -\frac{2}{N}\sum_{x=0}^{n_0} \left[x(Nq - n + x) - (n-x)(Np - x)\right] f(x)$$

Let, $g(x) = x(qN + x - n) f(x)$

$$= \frac{x(Nq - n + x)\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$$

$$= \frac{\frac{(Np)!}{(x-1)!\,(pN-x)!} \times \frac{(Nq)!}{(n-x)!\,(Nq+x-n-1)!}}{\binom{N}{n}}$$

$$g(x+1) = \frac{\frac{(Np)!}{x!\,(Np-x-1)!} \times \frac{(Nq)!}{(n-x-1)!\,(Nq+x-n)!}}{\binom{N}{n}}$$

$$= (Np - x)(n - x) f(x)$$

$$\therefore MD_\mu(X) = -\frac{2}{N}\left[\sum_{x=0}^{n_0}\{g(x) - g(x+1)\}\right]$$

$$= \frac{2}{N} g(n_0 + 1)$$

$$= \frac{2}{N} g([\mu] + 1)$$

$$= \frac{2}{N}([\mu]+1)\left[Nq - n + [\mu] + 1\right] f(x)$$

$$= \frac{2}{N}(Np - [\mu])(n - [\mu]) f(x).$$

☆ **Problem 1.** If $X \sim H(N, n, p)$ then find the distribution of $(n-X)$.

Soln →
$$P[X=x] = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$$

$$\therefore P[X=n-x] = \frac{\binom{Np}{n-x}\binom{Nq}{x}}{\binom{N}{n}}$$

$$\therefore P[n-X=x] = \frac{\binom{Np}{n-x}\binom{Nq}{x}}{\binom{N}{n}}$$

Let, $Y = n - X$.

$$\therefore P[Y=x] = \frac{\binom{Nq}{x}\binom{Np}{n-x}}{\binom{N}{n}}$$, where $Y \sim H(n, N, p)$.

☆ **Problem 2.** If $X \sim H(N, n, p)$ then show that $V(X) \le \frac{n}{4}$.

Soln. →
$$V(X) = \frac{npq(N-n)}{N-1}.$$

$$\frac{dV(X)}{dp} = \frac{n(N-n)}{N-1}[q-p] = 0$$

$$\therefore p = \frac{1}{2}.$$

$$\therefore \frac{d^2V(X)}{dp^2} = \frac{n(N-n)}{N-1}[-1-1] < 0$$

$\therefore V(X)$ is maximum at $p = \frac{1}{2}$.

$$\therefore V(X) \le \frac{n}{4} \cdot \frac{N-n}{N-1},$$ since $\frac{N-n}{N-1} \le 1$ as $n \geq 1$

$$\therefore Var(X) \le \frac{n}{4}.$$

☆ Problem 3. If $X \sim H(N, n, M)$ then show that —

$$E(X^r) = \frac{nM}{N} E(1+Y)^{r-1} \text{ where } Y \sim H(N-1, n-1, M-1).$$

Soln. →

$$E(X^r) = \sum_{x=0}^{n} x^r \cdot \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=0}^{n} x^{r-1} \frac{x \cdot \frac{M!}{x!(M-x)!}}{\binom{N}{n}} \cdot \binom{N-M}{n-x}$$

$$= \frac{M}{\binom{N}{n}} \sum_{x=1}^{n} x^{r-1} \binom{M-1}{x-1}\binom{N-M}{n-x} \qquad \text{Put, } x-1 = x'$$

$$= \frac{nM}{N\binom{N-1}{n-1}} \sum_{x'=0}^{n-1} (x'+1)^{r-1} \binom{M-1}{x'}\binom{N-M}{n-x'-1}$$

$$= \frac{nM}{N} \sum_{x'=0}^{n-1} (x'+1)^{r-1} \cdot \frac{\binom{M-1}{x'}\binom{N-M}{n-1-x'}}{\binom{N-1}{n-1}}$$

$$= \frac{nM}{N} E(1+Y)^{r-1}, \text{ where } Y \sim H(N-1, n-1, M-1).$$

☆ Problem 4. From a lot of N objects; of which Np are of a given kind, objects are drawn one by one without replacement till r objects of the given kind are obtained. Find the probability distribution of the no. of drawings required (say X). S.T. —

$$E(X) = r \times \frac{N+1}{Np+1}, \quad Var(X) = r \times \frac{Nq}{Np+1} \times \frac{Np+1-r}{Np+1} \times \frac{N+1}{Np+2},$$

where $q = 1-p$.

(This is called an inverse hypergeometric distribution)

Soln. → Mass points of X are $r, r+1, \ldots, r+Nq$

if x be any such mass point then,

$P[X=x] = P$[exactly x drawings are required to get 'r' objects of the given kind]

$= P$[in the first $x-1$ drawing $r-1$ objects of the given kind and $x-r$ objects of the other kinds are obtained] $\times$ P[an object of the given kind came in the last trial]

$$= \frac{\binom{Np}{r-1}\binom{Nq}{x-r}}{\binom{N}{x-1}} \cdot \frac{Np-r+1}{N-x+1}$$

$$= \frac{\dfrac{Np!\,(Np-r+1)}{(r-1)!\,(Np-r+1)!} \times \dfrac{Nq!}{(x-r)!\,(Nq-x+r)!}}{\dfrac{N!}{(x-1)!\,(N-x+1)!} \times (N-x+1)}$$

$$= \frac{\dfrac{Np!}{(r-1)!\,(Np-r)!} \times \dfrac{Nq!}{(x-r)!\,(Nq-x+r)!}}{\dfrac{N!}{(x-1)!\,(N-x)!}}$$

$$= \frac{\dfrac{(x-1)!}{(r-1)!\,(x-r)!} \times \dfrac{(N-x)!}{(Np-r)!\,(N-Np-x+r)!}}{\dfrac{N!}{(N-Np)!\,Np!}}$$

$$= \binom{x-1}{r-1} \cdot \frac{\binom{N-x}{Np-r}}{\binom{N}{Np}}$$

$$E(X) = \sum_{x=r}^{r+Nq} x \binom{x-1}{r-1} \cdot \frac{\binom{N-x}{Np-r}}{\binom{N}{Np}}$$

$$= r \sum_{x=r}^{Nq+r} \binom{x}{r} \frac{\binom{N-x}{Np-r}}{\binom{N}{Np}}$$

Let, $x' = x+1$, $r' = r+1$, $(Np)' = Np+1$, $N' = N+1$, $(Nq)' = N' - (Np)'$

$$= r \sum_{x'=r'}^{(Nq)'+r'} \binom{x'-1}{r'-1} \frac{\binom{N'-x'}{(Np)'-r'}}{\binom{N}{Np}}$$

$$= \frac{r}{\binom{N}{Np}} \binom{N+1}{Np+1}$$

$$= \frac{r\,(N+1)}{(Np+1)}.$$

$$E[X(X+1)] = \sum_{x=r}^{r+Nq} x(x+1)\binom{x-1}{r-1}\cdot\frac{\binom{N-x}{Np-r}}{\binom{N}{Np}}$$

$$= r(r+1)\sum_{x=r}^{Nq+r}\binom{x+1}{r+1}\cdot\frac{\binom{N-x}{Np-r}}{\binom{N}{Np}}$$

Let, $x'=x+2$, $r'=r+2$, $(Np)'=Np+2$, $N'=N+2$, $(Nq)'=N'-(Np)'$

$$= r(r+1)\sum_{x'=r'}^{(Nq)'+r'}\binom{x'-1}{r'-1}\frac{\binom{N'-x'}{(Np)'-r}}{\binom{N}{Np}}$$

$$= \frac{r(r+1)}{\binom{N}{Np}}\binom{N+2}{Np+2}$$

$$= \frac{r(r+1)(N+1)(N+2)}{(Np+2)(Np+1)}$$

$$\therefore Var(X) = \frac{r(r+1)\,(N+1)(N+2)}{(Np+1)(Np+2)} - \frac{r(N+1)}{(Np+1)} - \frac{r^2(N+1)^2}{(Np+1)^2}$$

$$= \frac{r(N+1)}{(Np+1)^2(Np+2)}\Big[\cancel{rN^2p} + \cancel{2rNp} + N^2p + \cancel{2Np} + \cancel{rN} + \cancel{2r} + N + \cancel{2} - N^2p^2 - \cancel{3}Np - \cancel{2} - \cancel{rN^2p} - \cancel{2Nr} - rNp - \cancel{2r}\Big]$$

$$= \frac{r(N+1)}{(Np+1)^2(Np+2)}\left[rNp + N^2p + N - N^2p^2 - Np - Nr\right]$$

$$= \frac{r(N+1)}{(Np+1)^2(Np+2)}\left[N^2pq + Nq - rNq\right]$$

$$= \frac{r(N+1)Nq}{(Np+1)^2(Np+2)}(Np+1-r)$$

Now,

if Y = # objects of the other kind preceeding the rth object of the given kind.

then $X = Y + r$

Mass points of Y are $0, 1, \ldots, Nq$.

$$P[Y=y] = P[X-r=y] = P[X=r+y] = \binom{r+y-1}{r-1}\frac{\binom{N-r-y}{Np-r}}{\binom{N}{Np}}$$

$$E(Y) = E(X) - r$$

$$= \frac{r(N+1)}{Np+1} - r = \frac{rNq}{Np+1}.$$

☆ **Problem 6.** If x be a mass point of a random variable, having hypergeometric distribution with parameters (N, n, p); show that, —

$$\binom{n}{x}\left(p-\frac{x}{N}\right)^{x}\left(q-\frac{n-x}{N}\right)^{n-x} < P[X=x] < \binom{n}{x}p^{x}q^{n-x}\left(1-\frac{n}{N}\right)^{-n}$$

Hence justify the approximation of $P[X=x]$ by the binomial probability, $\binom{n}{x}p^{x}q^{n-x}$.

Soln. → $\boxed{}$ $P[X=x] = \dfrac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}}$, for a particular x

$$= \binom{n}{x}\frac{(Np)_x\,(Nq)_{n-x}}{(N)_n}$$

$$= \binom{n}{x}\frac{Np(Np-1)\cdots(Np-x+1)\,Nq(Nq-1)\cdots(Nq-n+x+1)}{N(N-1)\cdots\cdots(N-n+1)}$$

$$= \binom{n}{x}\frac{\cancel{N^{x}}\cdot\cancel{N^{n-x}}\cdot p\cdot\left(p-\frac{1}{N}\right)\left(p-\frac{2}{N}\right)\cdots\left(p-\frac{x-1}{N}\right)q\left(q-\frac{1}{N}\right)\cdots\left(q-\frac{n-x-1}{N}\right)}{\cancel{N^{n}}\cdot 1\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\cdots\left(1-\frac{n-1}{N}\right)}$$

$= \binom{n}{x} ABC$, say, where,

$$A = p\left(p-\frac{1}{N}\right)\cdots\left(p-\frac{x-1}{N}\right).$$

$$B = q\left(q-\frac{1}{N}\right)\cdots\cdots\left(q-\frac{n-x-1}{N}\right).$$

$$C = \frac{1}{\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)}.$$

R.H.S

$p-\frac{1}{N}<p$ | $q-\frac{1}{N}<q$

$p-\frac{2}{N}<p$ | $q-\frac{2}{N}<q$

$\vdots$ | $\vdots$

$p-\frac{x-1}{N}<p$ | $q-\frac{n-x-1}{N}<q$

$\Rightarrow \frac{A}{p}<p^{x-1}$ | $\Rightarrow \frac{B}{q}<q^{n-x-1}$

$\therefore A<p^{x}$. | $\therefore B<q^{n-x}$.

$n > 1$

$\Rightarrow \frac{1}{N} < \frac{n}{N}$

$\Rightarrow 1 - \frac{1}{N} > 1 - \frac{n}{N}$

$1 - \frac{2}{N} > 1 - \frac{n}{N}$

$\vdots$

$1 - \frac{n-1}{N} > 1 - \frac{n}{N}$

$\Rightarrow \frac{1}{C} > \left(1 - \frac{n}{N}\right)^n$

$\therefore C < \left(1 - \frac{n}{N}\right)^{-n}$

$\therefore P[X=x] = \binom{n}{x} ABC < \binom{n}{x} p^x q^{n-x} \left(1 - \frac{n}{N}\right)^{-n}$

L.H.S

$p > p - \frac{x}{N}$ | $q > q - \frac{n-x}{N}$

$p - \frac{1}{N} > p - \frac{x}{N}$ | $q - \frac{1}{N} > q - \frac{n-x}{N}$

$\vdots$

$p - \frac{x-1}{N} > p - \frac{x}{N}$ | $q - \frac{n-x-1}{N} > q - \frac{n-x}{N}$

$\Rightarrow A > \left(p - \frac{x}{N}\right)^x$ | $\Rightarrow B > \left(q - \frac{n-x}{N}\right)^{n-x}$

$1 > 1 - \frac{1}{N}$

$1 > 1 - \frac{2}{N}$

$\vdots$

$1 > 1 - \frac{n-1}{N}$

$\Rightarrow C < 1$

$\Rightarrow \frac{1}{C} > 1$

$\therefore P[X=x] = \binom{n}{x} ABC > \binom{n}{x} \left(p - \frac{x}{N}\right)^x \left(q - \frac{n-x}{N}\right)^{n-x}$

Combining L.H.S & R.H.S, we get the required result.

Now, $\lim_{N \to \infty} \binom{n}{x} \left(p - \frac{x}{N}\right)^x \left(q - \frac{n-x}{N}\right)^{n-x} = \binom{n}{x} p^x q^{n-x}$

and

$\lim_{n \to \infty} \binom{n}{x} p^x q^{n-x} \cdot \left(1 - \frac{n}{N}\right)^{-n} = \binom{n}{x} p^x q^{n-x}$

Hence by squeeze or sandwich theorem,

$\lim_{n \to \infty} P[X=x] = \binom{n}{x} p^x q^{n-x}$.

DISCRETE UNIFORM DISTRIBUTION :→

(65)

Let X be a r.v. takes only the values $x_1, \ldots, x_N$. Then X is said to have a uniform distribution over the set $\{x_1, x_2, \ldots, x_N\}$ if $P[X = x_i] = \text{constant } (K)$, $i = 1(1)N$.

Now, $$1 = P[X = \Omega] = P\left[\bigcup_{i=1}^{N}\{X = x_i\}\right]$$
$$= \sum_{i=1}^{N} P[X = x_i]$$
$$= \sum_{i=1}^{N} K$$
$$= NK.$$

$$\Rightarrow K = \frac{1}{N}.$$

- **Definition:—** A R.V. X is said to have a uniform distribution over $\{x_1, x_2, \ldots, x_N\}$ if its PMF is —

$$f_X(x) = \begin{cases} \frac{1}{N} & \text{if } x = x_i \quad \forall\, i = 1(1)N \\ 0 & \text{o/w} \end{cases}$$

- **MOMENTS:—**

$$E(X) = \sum_{i=1}^{N} x_i P[X = x_i]$$
$$= \frac{1}{N}\sum_{i=1}^{n} x_i$$
$$= \bar{x}, \text{ (say)}$$

$$Var(X) = E(X^2) - E^2(X)$$
$$= \frac{1}{N}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$$
$$= \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 = s^2, \text{ (say)}.$$

In descriptive statistics, if we assume that the sampling values are uniformly distributed then,

$$\text{Sample Mean} = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \text{ and}$$

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \text{Variance}.$$

POWER SERIES DISTRIBUTION : $\rightarrow$

Definition: — A discrete random variable X with mass points 0, 1, 2, ... is said to follow Power series Distribution if its p.m.f. is of the following form, —

$$f(x) = \frac{a_x \cdot \theta^x}{g(\theta)} \quad , x = 0, 1, 2, \ldots$$

where,

$g(\theta) = \sum_{x=0}^{\infty} a_x \theta^x$ & $0 < \theta < r$ [r is the radious of convergence of the convergent Power series]

$\theta > 0$ and $g(\theta)$ is differentiable & a_x is non negative.

Example: —

1) $X \sim Bin(n, p)$

Let us take, $a_x = \binom{n}{x}$, $x = 0, 1, \ldots, n$

$$\therefore g(\theta) = \sum_{x=0}^{\infty} \binom{n}{x} \theta^x = (1+\theta)^n$$

$$\therefore f(x) = \binom{n}{x} \frac{\theta^x}{(1+\theta)^n} = \binom{n}{x} \left(\frac{\theta}{1+\theta}\right)^x \left(1 - \frac{\theta}{1+\theta}\right)^{n-x}$$

$$= \binom{n}{x} p^x q^{n-x} \quad ; \quad p + q = 1, \quad 0 < p < 1$$

where, $p = \frac{\theta}{1+\theta}$, $q = \frac{1}{1+\theta}$; $\theta > 0$

2) $X \sim P(\lambda)$

Let us take, $a_x = \frac{1}{x!}$, $x = 0, 1, 2, \ldots$

$$g(\theta) = \sum_{x=0}^{\infty} \frac{\theta^x}{x!} = e^{\theta}$$

$$\therefore f(x) = \frac{1}{x!} \cdot \theta^x \cdot e^{-\theta} = \frac{e^{-\theta} \cdot \theta^x}{x!} \quad , \theta > 0$$

$$= \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad ; \quad \theta = \lambda > 0$$

3) $X \sim NB(r, p)$

Let us take, $a_x = \binom{r+x-1}{x}$, $x = 0, 1, 2, \ldots$; $r > 0$

$$g(\theta) = \sum_{x=0}^{\infty} \binom{r+x-1}{x} \theta^x = (1-\theta)^{-r} \quad , 0 < \theta < 1$$

$$\therefore f(x) = \binom{r+x-1}{x} \cdot \frac{\theta^x}{(1-\theta)^{-r}}$$

$$= \binom{x+r-1}{x} \theta^x (1-\theta)^r \quad \left[(1-\theta) = p , \theta = q \right]$$

$$= \binom{x+r-1}{x} q^x p^r$$

4) Let us take, $a_x = 1$, $\forall x = 0, 1, 2, \dots$

$$\therefore g(\theta) = \sum_{x=0}^{\infty} \theta^x = (1-\theta)^{-1} \quad ; \because 0 < \theta < 1$$

$$\therefore f(x) = (1-\theta) \cdot \theta^x$$

$$= pq^x \quad ; \; p = (1-\theta), \; q = \theta$$

$$\therefore \underline{X \sim G(p)}$$

5) Let us take, $a_x = \frac{1}{x}$, $x = 1, 2, \dots$

$= 0$ OW

$$g(x) = \sum_{x=1}^{\infty} \frac{1}{x} \cdot \theta^x$$

$$\therefore g(x) = -\log(1-\theta)$$

$$\therefore f(x) = -\frac{1}{x} \cdot \frac{\theta^x}{\log(1-\theta)}$$

$$\therefore \underline{X \sim \text{Logarithmic}(\theta)}$$

P.m.f. As we know, $\displaystyle\sum_{x=0}^{\infty} \frac{a_x \cdot \theta^x}{g(\theta)} = 1$

, i.e. $g(\theta) = \displaystyle\sum_{x=0}^{\infty} a_x \theta^x$.

■ **Probability Generating Function:** $\longrightarrow$ Let $X \sim$ Power series Distn.

$$\text{PGF} = P_X(t) = E(t^X)$$

$$= \sum_{x=0}^{\infty} t^x \cdot \frac{a_x \cdot \theta^x}{g(\theta)} = \sum_{x=0}^{\infty} \frac{a_x \cdot (\theta t)^x}{g(\theta)}$$

$$= \frac{g(t\theta)}{g(\theta)}$$

$$\therefore E(X) = P'(1) = \frac{d}{dt}\left\{\frac{g(\theta t)}{g(\theta)}\right\}_{t=1}$$

$$= \frac{\theta g'(\theta)}{g(\theta)} = \theta \cdot \frac{d}{d\theta} \ln g(\theta)$$

$$P''(1) = E[X(X-1)]$$

$$= \frac{d^2}{dt^2}\left\{\frac{g(\theta t)}{g(\theta)}\right\}_{t=1}$$

$$= \frac{\theta^2 g''(\theta)}{g(\theta)}$$

$$\therefore V(X) = E\left(X(X-1)\right) + E(X) - E^2(X)$$

$$= \frac{\theta^2 g''(\theta)}{g(\theta)} + \frac{\theta g'(\theta)}{g(\theta)} - \left[\frac{\theta g'(\theta)}{g(\theta)}\right]^2$$

$$= \theta^2\left[\frac{g(\theta)g''(\theta) - \{g'(\theta)\}^2}{g^2(\theta)}\right] + \frac{\theta g'(\theta)}{g(\theta)}$$

$$= \theta^2 \cdot \frac{d}{d\theta}\left(\frac{g'(\theta)}{g(\theta)}\right) + \theta \cdot \frac{d}{d\theta}\ln g(\theta)$$

$$\therefore V(X) = \theta^2 \cdot \frac{d^2}{d\theta^2}\ln g(\theta) + \theta \cdot \frac{d}{d\theta}\ln g(\theta).$$

Example 1:-

i) If $X \sim Bin(n, p)$,

then $a_x = \binom{n}{x}$ and $\theta = \frac{p}{1-p} \Rightarrow p = \frac{\theta}{1+\theta}$

and $g(\theta) = (1+\theta)^n$

$$\therefore E(X) = \theta \cdot \frac{d}{d\theta}\ln(1+\theta)^n$$

$$= n\theta \cdot \frac{1}{1+\theta} = np$$

$$\therefore V(X) = \theta^2 \cdot \frac{d^2}{d\theta^2} \cdot \ln(1+\theta)^n + np = \theta^2 \cdot n \cdot \frac{(-1)}{(1+\theta)^2} + np = -np^2 + np = npq.$$

ii) If $X \sim P(\lambda)$.

then $a_x = \frac{1}{x!}$ and $\theta = \lambda$ and $g(\theta) = e^\theta$

$$\therefore E(X) = \theta \cdot \frac{d}{d\theta} \cdot \ln e^\theta = \theta = \lambda$$

$$\therefore V(X) = \theta^2 \cdot \frac{d^2}{d\theta^2}\ln e^\theta + \lambda$$

$$= 0 + \lambda = \lambda$$

iii) If $X \sim NB(r, p)$

Then, $a_x = \binom{x+r-1}{x}$, $\theta = (1-p) = q$

$$\therefore g(\theta) = p^{-r} = (1-q)^{-r} = (1-\theta)^{-r}$$

$$\therefore E(X) = \theta \cdot \frac{d}{d\theta}\ln(1-\theta)^{-r} = -\frac{r \cdot \theta \cdot (-1)}{1-\theta} = \frac{rq}{p}$$

$$\therefore V(X) = \theta^2 \cdot \frac{d^2}{d\theta^2}\ln(1-\theta)^{-r} + r \cdot \frac{q}{p}$$

$$= \theta^2 \cdot (-r) \cdot \frac{d}{d\theta}\left(-\frac{1}{1-\theta}\right) + r \cdot \frac{q}{p}$$

$$= \theta^2 \cdot r \cdot \frac{1}{(1-\theta)^2} + r \cdot \frac{q}{p}$$

$$= r\left(\frac{q^2}{p^2}\right) + r\left(\frac{q}{p}\right)$$

$$= \frac{rq}{p}\left(\frac{q}{p} + 1\right) = \frac{rq}{p^2}$$

Recursion Relation for Central moments :—

Let $\mu = f(\theta) = \frac{\theta g'(\theta)}{g(\theta)}$ = a function of θ = mean.

$$P[X=x] = \frac{a_x \theta^x}{g(\theta)}, \quad x = 0, 1, \ldots$$

Now,
$$\mu_r = E(X-\mu)^r$$
$$= \sum_{x=0}^{\infty} (x - f(\theta))^r \cdot \frac{a_x \cdot \theta^x}{g(\theta)}$$

Differentiating both sides w.r.t. θ, we get, —

$$\frac{d\mu_r}{d\theta} = \sum_{x=0}^{\infty} (-r)[x - f(\theta)]^{r-1} \cdot f'(\theta) \cdot \frac{a_x \cdot \theta^x}{g(\theta)} +$$
$$\sum_{x=0}^{\infty} [x - f(\theta)]^r \cdot a_x \left[\frac{x\theta^{x-1}}{g(\theta)} - \theta^x \cdot \frac{g'(\theta)}{g^2(\theta)}\right]$$

$$= (-r) f'(\theta) \sum_{x=0}^{\infty} [x - f(\theta)]^{r-1} \cdot \frac{a_x \cdot \theta^x}{g(\theta)} +$$
$$\sum_{x=0}^{\infty} [x - f(\theta)]^r \cdot \frac{a_x}{g(\theta)} \cdot \theta^{x-1}\left(x - \theta \cdot \frac{g'(\theta)}{g(\theta)}\right)$$

$$= (-r) f'(\theta)\, \mu_{r-1} + \sum_{x=0}^{\infty} (x - f(\theta))^{r+1} \frac{a_x}{g(\theta)} \cdot \theta^{x-1}$$

$$= (-r) f'(\theta) \mu_{r-1} + \frac{1}{\theta} \mu_{r+1}$$

$$\therefore \boxed{\mu_{r+1} = \theta\left[\frac{d\mu_r}{d\theta} + r f'(\theta)\, \mu_{r-1}\right]}$$

Note, Putting, $r=1$, $\mu_2 = \theta \cdot f'(\theta)$

Example:—

1) If $X \sim Bin(n, p)$

$\theta = \frac{p}{q}$; $p = \frac{\theta}{1+\theta}$, $\quad g(\theta) = (1+\theta)^n$ and $f(\theta)$ = mean = np

$$\mu_{r+1} = \frac{p}{q}\left[\frac{d\mu_r}{dp} \cdot \frac{dp}{d\theta} + r\left(\frac{d}{dp} \cdot np \cdot \frac{dp}{d\theta}\right) \cdot \mu_{r-1}\right]$$

$$= \frac{p}{q}\left[\frac{d\mu_r}{dp}\left(\frac{d}{d\theta} \cdot \frac{\theta}{1+\theta}\right) + nr\left(\frac{d}{d\theta} \cdot \frac{\theta}{1+\theta}\right) \cdot \mu_{r-1}\right]$$

$$= \frac{p}{q}\left[\frac{d\mu_r}{dp} \cdot \frac{1}{(1+\theta)^2} + nr \cdot \frac{1}{(1+\theta)^2} \cdot \mu_{r-1}\right]$$

$$= \frac{p}{q}\left[q^2 \cdot \frac{d\mu_r}{dp} + q^2 \cdot nr \cdot \mu_{r-1}\right]$$

$$= pq\left[\frac{d\mu_r}{dp} + nr\mu_{r-1}\right] \text{ for } r = 1, 2, \ldots$$

2) If $X \sim P(\lambda)$.

$\theta = \lambda, \quad \mu = f(\theta) = \lambda$

$g(\theta) = e^{\theta} = e^{\lambda}$

$$\mu_{r+1} = \lambda\left[\frac{d}{d\lambda}\cdot\mu_r + r\mu_{r-1}\right] \quad \text{for } r = 1, 2, \ldots$$

3) If $X \sim NB(r, p)$.

$\mu = f(\theta) = \frac{rq}{p}$

$\theta = (1-p) \; ; \quad g(\theta) = p^{-r}$

$$\mu_{k+1} = q\left[\frac{d\mu_k}{dp}\cdot\frac{dp}{d\theta} + \left(k\cdot\frac{d}{dp}\cdot\frac{r(1-p)}{p}\right)\cdot\frac{dp}{d\theta}\mu_{k-1}\right]$$

$$= q\left[-\frac{d\mu_k}{dp} + \frac{rk}{p^2}\mu_{k-1}\right] \text{ for } k = 1, 2, \ldots$$

Recursion Relation for Raw Moments : →

$$\mu_r' = E(X^r) = \sum_{x=0}^{\infty} x^r \cdot \frac{a_x \cdot \theta^x}{g(\theta)}$$

Differentiating both sides w.r.t. θ, we get, —

$$\frac{d\mu_r'}{d\theta} = \sum_{x=0}^{\infty} x^r \cdot a_x\left[x\cdot\frac{\theta^{x-1}}{g(\theta)} - \theta^x \cdot \frac{g'(\theta)}{g^2(\theta)}\right]$$

$$= \frac{1}{\theta}\sum_{x=0}^{\infty} x^{r+1} a_x \cdot \frac{\theta^x}{g(\theta)} - \frac{1}{\theta}\sum_{x=0}^{\infty} x^r \cdot a_x \cdot \frac{\theta^x}{g(\theta)}\cdot f(\theta)$$

$$= \frac{1}{\theta}\mu_{r+1}' - \frac{1}{\theta}\cdot\mu_r'\cdot f(\theta)$$

$$\Rightarrow \mu_{r+1}' = \theta\cdot\frac{d}{d\theta}\mu_r' + \mu_r'\cdot f(\theta) \quad \text{------} (*)$$

Putting $r = 1$,

$$\mu_2' = E(X^2) = \theta\cdot\frac{d}{d\theta}\cdot\{f(\theta)\} + [f(\theta)]^2$$

☆ **Problem** :→ Show that the equality of mean and variance is a necessary & sufficient condition for a power series distribution to be Poisson.

Proof:-

If Part :→ Let $X \sim P(\theta)$

∴ and Poisson distribution also belongs to the Power series family.

and then it is known that if $X \sim P(\theta)$ then

$$E(X) = V(X) = \theta.$$

Only if Part :→ $X \sim$ Power Series Distribution

and $E(X) = V(X)$

$\therefore \mu_2 = \mu_1'$

$\Rightarrow \theta \cdot \dfrac{d\mu_1'}{d\theta} = \mu_1'$

$\Rightarrow \dfrac{d\mu_1'}{d\theta} = \dfrac{\mu_1'}{\theta}$

$\Rightarrow \dfrac{d\mu_1'}{\mu_1'} = \dfrac{d\theta}{\theta}$

$\Rightarrow \ln\mu_1' = \ln\theta + \ln c$ [$\ln c$ = constant]

$\therefore \mu_1' = \theta c$

$\Rightarrow \cancel{\theta} \cdot \dfrac{d}{d\theta} \ln(g(\theta)) = \cancel{\theta} c$

$\Rightarrow d\ln g(\theta) = c\,d\theta$

$\Rightarrow \ln g(\theta) = c\theta + k_1$

$\Rightarrow g(\theta) = e^{k_1} \cdot e^{c\theta} = k e^{c\theta}$ [let, $k = e^{k_1}$]

$\Rightarrow \displaystyle\sum_{x=0}^{\infty} a_x \theta^x = k \sum_{x=0}^{\infty} \frac{(c\theta)^x}{x!}$

$\therefore a_x = \dfrac{k \cdot c^x}{x!}$

$$\therefore f(x) = \frac{\frac{k \cdot c^x}{x!} \cdot \theta^x}{\cancel{k} e^{c\theta}}$$

$$= \frac{(c\theta)^x \cdot e^{-c\theta}}{x!}$$

$$= \frac{\lambda^x \cdot e^{-\lambda}}{x!} \quad [\text{let, } c\theta = \lambda]$$

$\therefore X \sim P(\lambda)$.

LOGARITHMIC SERIES DISTRIBUTION :

Definition:- The probability distribution of the discrete Random variable X is said to be a Logarithmic distribution with parameter θ if its probability distribution has the p.m.f. of the form

$$f(x) = \begin{cases} \alpha \cdot \frac{\theta^x}{x} & \text{for } x = 1, 2, 3, \ldots \; ; \; 0<\theta<1 \\ 0 & \text{OW} \end{cases}$$

where $\alpha = -\frac{1}{\log(1-\theta)}$; $\therefore f(x) = \frac{\theta^x}{-x\log(1-\theta)}$; $0<\theta<1$

We know that, $-\log(1-\theta) = \theta + \frac{\theta^2}{2} + \frac{\theta^3}{3} + \cdots + \infty = \sum_{x=1}^{\infty} \frac{\theta^x}{x}$

Now, $\sum_{x=1}^{\infty} \frac{\theta^x}{-x\log(1-\theta)} = -\frac{1}{\log(1-\theta)} \sum_{x=1}^{\infty} \frac{\theta^x}{x}$

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

Recursion Relation for Probability :

$$\frac{P(X=x+1)}{P(X=x)} = \frac{x\theta}{(x+1)}.$$

$\therefore P(X=x+1) = \frac{x\theta}{(x+1)} \cdot P(X=x)$ $[\theta<1]$

Hence the maximum value of $P[X=x]$ is at the initial value $x=1$; and the value of $P[X=x]$ decreases as x increases.

MGF :-

$$M(t) = E(e^{tx}) = \frac{\log(1-\theta e^t)}{\log(1-\theta)}.$$

PGF :-

$$P(t) = E(t^X) = \frac{\log(1-\theta t)}{\log(1-\theta)}.$$

Factorial Moments :-

$$\mu_{[r]} = E(X^{r}) = \alpha\theta^r \sum_{x=r}^{\infty} (x-1)(x-2)\cdots(x-r+1)\cdot\theta^{x-r}$$

$$= \alpha \cdot \theta^r \cdot \frac{d^{r-1}}{d\theta^{r-1}} \left[\sum_{x=1}^{\infty} \theta^{x-1}\right]$$

$$= \alpha\theta^r \cdot (r-1)! \, (1-\theta)^{-r}$$

$$= \frac{\theta^r (r-1)!}{-\log(1-\theta)(1-\theta)^r}.$$

The first four moments about the origin are —

$$\mu_1' = \alpha\theta(1-\theta)^{-1}$$

$$\therefore \mu_1' = \frac{\theta}{-(1-\theta)\log(1-\theta)}$$

$$\mu_2' = \alpha\theta^2(1-\theta)^{-2} + \alpha\theta(1-\theta)^{-1}$$

$$\therefore \mu_2' = \alpha\theta(1-\theta)^{-2}.$$

$$\therefore \mu_3' = \alpha\theta(1+\theta)(1-\theta)^{-3}$$

$$\therefore \mu_4' = \alpha\theta(1+4\theta+\theta^2)(1-\theta)^{-4}.$$

Central moments are —

$$Var(X) = \alpha\theta(1-\alpha\theta)(1-\theta)^{-2}$$

$$\mu_3 = \alpha\theta(1+\theta-3\alpha\theta+2\alpha^2\theta^2)(1-\theta)^{-3}$$

$$\mu_4 = \alpha\theta\{1+4\theta+\theta^2-4\alpha\theta(1+\theta)+6\alpha^2\theta^2-3\alpha^3\theta^3\}(1-\theta)^{-4}.$$

The moment ratios are —

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \beta_2 = \frac{\mu_4}{\mu_2^2}.$$

Both this tends to ∞ as θ tends to 0 or as θ tends to 1, with

$$\lim_{\theta\to 0}\left(\frac{\beta_2}{\beta_1}\right) = 1; \quad \lim_{\theta\to 1}\left(\frac{\beta_2}{\beta_1}\right) = \frac{3}{2}.$$

NOTE:— Logarithmic series is a power series distn. and Logarithmic series distribution is an increasing function of θ.

Problem 1.

☆ Calculation of mean: → $\frac{f(x+1)}{f(x)} = \frac{\theta x}{(x+1)}.$

$$\Rightarrow (x+1)f(x+1) = \theta x f(x) \quad \cdots\cdots (1)$$

$$\Rightarrow \sum_{x=1}^{\infty}(x+1)f(x+1) = \theta\sum_{x=1}^{\infty} x f(x)$$

$$\Rightarrow \sum_{y=2}^{\infty} y f(x) = \theta E(X) \quad [\text{Taking } x+1 = y]$$

$$\Rightarrow \sum_{y=1}^{\infty} y f(x) - \frac{\theta}{(-1)\log(1-\theta)} = \theta E(X)$$

$$\Rightarrow E(X) - \theta E(X) = -\frac{\theta}{\log(1-\theta)}$$

$$\Rightarrow E(X) = -\frac{\theta}{(1-\theta)\log(1-\theta)}.$$

Problem 2.

☆ **Calculation of variance :—**

Multiplying (1) by x, we get —

$$\theta x^2 f(x) = x^2 f(x+1) + x f(x+1)$$

$$\Rightarrow \theta \sum_{x=1}^{\infty} x^2 f(x) = \sum_{x=1}^{\infty} x^2 f(x+1) + \sum_{x=1}^{\infty} x f(x+1)$$

$$\Rightarrow \theta E(X^2) = \sum_{x'=2}^{\infty} (x'-1)^2 f(x') + \sum_{x'=2}^{\infty} (x'-1) f(x')$$

$$\Rightarrow \theta E(X^2) = \sum_{x'=1}^{\infty} (x'-1)^2 f(x') + \sum_{x'=1}^{\infty} (x'-1) f(x')$$

$$\Rightarrow \theta E(X^2) = E(X'-1)^2 + E(X'-1)$$

$$\Rightarrow E(X^2)\theta = E(X)^2 + E(X)$$

$$\Rightarrow E(X^2) = \frac{E(X)}{(1-\theta)} = -\frac{\theta}{(1-\theta)^2 \log(1-\theta)}$$

$$\therefore V(X) = E(X^2) - E^2(X)$$

$$= \frac{\alpha\theta}{(1-\theta)^2} - \frac{\alpha^2\theta^2}{(1-\theta)^2} = \frac{\alpha\theta}{(1-\theta)^2}(1-\alpha\theta) = \mu\left(\frac{1}{1-\theta} - \mu\right).$$

☆ **General Problem :→** A random variable X has p.m.f.

$$f(x) = \frac{\binom{n}{x} \beta(a+x, n+b-x)}{\beta(a,b)} \ ; \ x = 0, 1, \ldots, n.$$

Show that → $E(X) = \frac{na}{a+b}$ & $V(X) = \frac{nab(a+b+n)}{(a+b)^2(a+b+1)}$.

[1997]

Soln. → The rth order factorial moment of X is given by, —

$$\mu_{[r]} = E[(X)_r]$$

$$= \sum_{x=0}^{n} (x)_r \binom{n}{x} \frac{\beta(a+x, n+b-x)}{\beta(a,b)}$$

$$= \frac{1}{\beta(a,b)} \sum_{x=0}^{n} (x)_r \binom{n}{x} \int_0^1 t^{a+x-1} (1-t)^{n+b-x-1} dt$$

$$= \frac{1}{\beta(a,b)} \int_0^1 t^{a-1} (1-t)^{b-1} \left[\sum_{x=0}^{n} (x)_r \binom{n}{x} t^x (1-t)^{n-x} \right] dt$$

[interchanging the order of summation and integration]

$$= \frac{1}{\beta(a,b)} \int_0^1 t^{a-1} (1-t)^{b-1} \cdot (n)_r \cdot t^r \, dt$$

[$E[(X)_r] = (n)_r \cdot p^r$, to be shown in Binomial distribution]

$$= \frac{(n)_r}{\beta(a,b)} \int_0^1 t^{a+r-1} (1-t)^{b-1} dt$$

$$= \frac{(n)_r \, \beta(a+r, b)}{\beta(a,b)}.$$

Putting, $r=1$,

$$\mu_{[1]} = \frac{\beta(a+1, b)}{\beta(a, b)} \cdot n$$

$$= \frac{na}{a+b}.$$

Putting, $r=2$,

$$\mu_{[2]} = E[X(X-1)]$$

$$= \frac{n(n-1)\beta(a+2, b)}{\beta(a, b)}$$

$$= \frac{n(n-1)\, a(a+1)}{(a+b)(a+b+1)}$$

$$\therefore Var(X) = \frac{n(n-1)\, a(a+1)}{(a+b)(a+b+1)} + \frac{na}{a+b} - \frac{n^2a^2}{(a+b)^2}$$

$$= \frac{na}{a+b}\left[\frac{(n-1)(a+1)}{(a+b+1)} - \frac{na}{a+b} + 1\right]$$

$$= \frac{na}{a+b}\left[\frac{-\cancel{a^2} - \cancel{a} + nb - ab - \cancel{b} + \cancel{a^2} + 2ab + b^2 + \cancel{a} + \cancel{b}}{(a+b+1)(a+b)}\right]$$

$$= \frac{nab\,(a+b+n)}{(a+b)^2(a+b+1)}. \quad \text{[Proved]}$$