$X \sim \text{Bin}(n, p)$.

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (p \cdot e^t)^x (1-p)^{n-x}$$

$$= (q + p \cdot e^t)^n \text{, where } q = 1-p$$

$$E(X) = \frac{d}{dt} M_X(t) \Big|_{t=0} = n (q + p \cdot e^t)^{n-1} p \cdot e^t \Big|_{t=0}$$

$$= np.$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0} = \frac{d}{dt} \left[n (q + p \cdot e^t)^{n-1} p \cdot e^t \right] \Big|_{t=0}$$

$$\overset{\text{algebra}}{=} npq + (np)^2$$

Then $\text{Var}(X) = E(X^2) - \{E(X)\}^2 = npq + (np)^2 - (np)^2 = npq$

An observation:- $Var(X) = npq \le np = E(X)$ since $0 \le q = 1-p \le 1$

This implies that for a given data, if we observe sample variance > Measure of central tendency, we can rull out the option of fitting by Binomial $dist^n$.

A real life example:- (Medical science problem Tay-Sachs).

Name of the disease

If A couple cary this genetic disordering, the prob. of having this genetic disording among their child is $= 0.25$.

Suppose A couple has 4 children.

X : No. of children has this gentic disordering.

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

$$P[X = x] = \binom{4}{x} (0.25)^x (1 - 0.25)^{4-x}.$$

Prob. computation:-

x	$P[X = x]$
0	0.316
1	0.422
2	0.211
3	0.047
4	0.004

$E(X) = 4 \times 0.25 = 1.$

Note that mode $(=1)$ = mean
$E(X) = 1$

[you can investigate the behaviours of median as well, and the ordering with mean & mode].

5) Negative binomial $dist^n$:-

Experiment:- Repeat indep. Bernoulli trials until having r success.

X: Number of failures preceding the r-th success.

Range space $\mathcal{X} = \{0, 1, 2, \cdots\cdots\}$.

p.m.f :-

$$p[X = x] = \begin{cases} 0 & \text{if } x \notin \mathcal{X} \\ \binom{x+r-1}{x} p^{r-1} (1-p)^x \times p & x \in \mathcal{X} \end{cases}$$

$$= \binom{x+r-1}{x} p^r (1-p)^x \quad \text{if } x \in \mathcal{X}$$

$X \sim NB(r, p)$.

i) Why $p[X = x]$ is a p.m.f ??

Note $p[X = x] \geq 0 \; \forall \, x, \; \forall \; p \in [0,1]$.

Next, $\sum_{x=0}^{\infty} p[X = x] = p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} q^x$, where $q = 1-p$.

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

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

check

$$= p^r (1-q)^{-r} = 1 \text{ since } q = 1-p.$$

$$\boxed{= (-1)^x \binom{-r}{x}}$$

check

observation

$$\binom{x+r-1}{x} = \frac{(x+r-1)(x+r-2)\cdots 1}{x!\,(r-1)!}$$

$$= \frac{(x+r-1)(x+r-2)\cdots r}{x!}$$

$$= \frac{(-1)^x \{(-r)(-r-1)\cdots(-x-r+1)\}}{x!}$$

ii) $M_X(t) = \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{x} p^r q^x$

MGF of X

$= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} (q \cdot e^t)^x$ ⟶ Using earlier argument.

$= p^r \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} (q \cdot e^t)^x$

$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-q \cdot e^t)^x$

$= p^r (1 - qe^t)^{-r}$

iii) $E(X) = \frac{d}{dt} M_X(t) \Big|_{t=0} = \frac{d}{dt} \left[p^r (1-qe^t)^{-r} \right] \Big|_{t=0}$

$= p^r(-r)(1-q \cdot e^t)^{-r-1}(-q \cdot e^t) \Big|_{t=0} = \frac{rq}{p}$.

iii) $E(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0} =$

Other option!—

$$E(X^2) = E[\underbrace{X(X-1)}_{\text{Not difficult.}}] + E(X)$$

We have $E[X(X-1)] \overset{\text{check}}{=} \frac{r(r+1)q^2}{p^2}$

Then $Var(X) = \underbrace{E[X(X-1)] + E(X)}_{E(X^2)} - \{E(X)\}^2$

$$= \frac{r(r+1)q^2}{p^2} + \frac{rq}{p} - \left\{\frac{rq}{p}\right\}^2$$

$$= \frac{rq}{p^2}.$$ Recall that $E(X) = \frac{rq}{p}$.

and hence, $Var(X) = \frac{rq}{p^2} > \frac{rq}{p} = E(X)$.

Remark:– For a given data, if spreadness $<$ the measure of central tendency, we can rule out the possibility of having negative binomial dist^n as the parent dist^n.

6) **Geometric dist^n:–**

Sp. case of $NB(r, p)$ when $r = 1$.

p.m.f.:– $$P[X = x] = p \cdot q^x \; ; \quad x = 0, 1, 2 \ldots$$

i) $P[X = x]$ is a p.m.f. (already we have done for general r).

A very special property of geometric dist^n

Lack of Memory property:– It states that
$$P[X > m+n \mid X > m] = P[X \geq n].$$

For $m, n \geq 0$,

$$P[X > m+n \mid X > m]$$

$$= \frac{P[X > m+n \cap X > m]}{P[X > m]}$$

$$= \frac{P[X > m+n]}{P[X > m]} = \frac{q^{m+n}}{q^m} = q^n = P[X \geq n]$$

✓

It proves the result.

[Check that $P[X \geq m] = q^m$.]

Using this