Special Discrete Distribution.

The Discrete uniform Distribution:

Let N: a given positive integer

$x_1 < x_2 < \cdots < x_N$ given real numbers

A r.v. X is said to follow a discrete uniform distribution on the set $\{x_1, \ldots, x_N\}$ written as $X \sim U(\{x_1, x_2, \ldots, x_N\})$ if its p.m.f. is given by

$$f_X(x) = P(X=x) = \begin{cases} \frac{1}{N}, & x \in \{x_1, \ldots, x_N\} \\ 0, & o/\omega. \end{cases}$$

In particular suppose $X \sim U(\{1, 2, \ldots, N\})$

$$\mu_1' = \mu = E(X) = \frac{1}{N}\sum_{i=1}^{N} i = \frac{(N+1)}{2}$$

$$\mu_2' = E(X^2) = \frac{1}{N}\sum_{i=1}^{n} i^2 = \frac{(N+1)(2N+1)}{6}$$

$$Var(X) = E(X^2) - (E(X))^2 = \frac{N^2-1}{12}$$

$$M_X(t) = E(e^{tX}) = \sum_{j=1}^{N} e^{tj} \cdot \frac{1}{N} = \begin{cases} \frac{e^t(e^{Nt}-1)}{N(e^t-1)}, & t \neq 0 \\ 1, & t = 0 \end{cases}$$

μ_K', $K = 1, 2, \ldots$ exists for all +ve integral value K.

Bernoulli Distribution:

Defn: A Bernoullian trial is an expt with two possible outcomes say success (S) and failure (F)

Let X be a random variable takes value 1 for S and 0 for F. The p.m.f

$$f_X(1) = P(X=1) = p.$$
$$f_X(0) = P(X=0) = (1-p) = q$$

$$E(X) = 0\cdot(1-p) + 1\cdot p = p$$

$$\mu_k' = E(X^k) = p, \quad k = 1,2,\ldots$$

$$V(X) = E(X^2) - (E(X))^2 = p(1-p) = pq$$

$$M_X(t) = E(e^{tX}) = (1-p)e^{t\cdot 0} + pe^t$$
$$= (1-p+pe^t)$$
$$= (q+pe^t).$$

Binomial Distribution:

Consider a sequence of n independent Bernoulli trials with prob of S in each trial as $p \in (0,1)$, $n \in \mathbb{N}$ is a fixed natural no.

Define X = number of success in n trials

$\rightarrow 0,1,2,\ldots n.$

$$S = \{0,1,2,\ldots n\}.$$

Then

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots n.$$

Define

$$f_X(x) = P(X=x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \dots n. \\ 0 & \text{o/w} \end{cases}$$

→ Binomial distⁿ with n trials and success probability p denoted by $Bin(n, p)$ and written as

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

$\{Bin(n, p) : n \in \mathbb{N}, p \in (0,1)\}$ is the family of prob. distributions has two parameter $n \in \mathbb{N}$ and $p \in (0,1)$.

M.g.f. $$M_X(t) = (1 - p + pe^t)^n, \quad t \in \mathbb{R}$$
$$= (q + pe^t)^n, \quad q = (1-p).$$

$E(X) = np$. $\quad Var(X) = npq$

$$\mu_3 = E(X - np)^3 = np(1-p)(1-2p) = npq(1-2p)$$

$$\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{npq(1-2p)}{(npq)^{3/2}} = \begin{cases} 0 & \text{if } p = 1/2 \text{ (symmetric)} \\ >0 & \text{if } p < 1/2 \text{ +vely skewed} \\ <0 & \text{if } p > 1/2 \text{ -vely skewed.} \end{cases}$$

Alternative Method :

Let $x^{(r)} = x(x-1)(x-2)\cdots(x-r+1)$

$E(x^{(r)})$ is called factorial moment

$$E(x^{(r)}) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} x(x-1)(x-2)\cdots(x-r+1)$$

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

$$= n(n-1)\cdots(n-r+1)\, p^r \sum_{x=r}^{n} \frac{(n-r)!\; p^{x-r} (1-p)^{n-x}}{(x-r)!\,(n-x)!}$$

take $x - r = i$

$$E(x^{(r)}) = n(n-1)\cdots(n-r+1)\, p^r \sum_{i=0}^{n-r} \frac{(n-r)!\; p^i (1-p)^{n-r-i}}{i!\,(n-r-i)!}$$

$$= n(n-1)\cdots(n-r+1)\, p^r (1-p+p)^{n-r} = n(n-1)\cdots(n-r+1)p^r.$$

we have

$$\mu_4 = E(x-np)^4 = pn(\cancel{1}-p)\left[3p^2(2-n) + 3p(n-2) + 1\right]$$

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

Example: A fair dice is rolled 5 times independently. Find the probability that on 3 occeassions we get a six.

Solun: Consider getting a six as success. Then

$X = \#$ of success in 5 trials

$\sim Bin(5, \frac{1}{6})$

Required probability $= P(X=3) = \binom{5}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$.

Geometric Distribution:

Independent Bernoullian trials aree performed till a success is achived. Let X denote the no. of trials needed to achive the first success.

$X \to 1, 2, 3, \ldots$

$P(X=j) = q^{j-1}p, \quad j=1,2,\ldots$

so the p.m.f is

$$f_X(x) = P(X=x) = q^{x-1}p, \quad x=1,2,\ldots$$

$$\sum_{x=1}^{\infty} f_X(x) = \sum_{x=1}^{\infty} P(X=x) = \sum_{x=1}^{\infty} q^{x-1}p = p(1+q+q^2+\ldots)$$

$$= \frac{p}{1-q} = 1.$$

$$\mu = E(x) = \sum_{x=1}^{\infty} x q^{x-1} p = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}.$$

$$\mu_2' = E(x^2) = \frac{q+1}{p^2}, \quad Var(X) = \sigma^2 = \frac{q+1}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

$$M_X(t) = \sum_{x=1}^{\infty} e^{tx} q^{x-1} p = pe^t \sum_{x=1}^{\infty} (qe^t)^{x-1}$$

$$= \frac{pe^t}{1-qe^t}, \quad 0 < qe^t < 1$$

In this case we denote $X \sim Geo(p)$, $p \in (0,1)$

Family is $\{Geo(p) : p \in (0,1)\}$. — one parameter.

Consider

$$P(X > m) = \sum_{x=m+1}^{\infty} q^{x-1} p = q^m p(1+q+\cdots) = q^m$$

$$P(\underbrace{X > m+n}_{E_1} \mid \underbrace{X > n}_{E_2}) = \frac{P(E_1 \cap E_2)}{P(E_2)} = \frac{P(E_1)}{P(E_2)}$$

$$= \frac{q^{m+n}}{q^n} = q^m = P(X > m)$$

So $P(X > m+n \mid X > n) = P(X > m)$ — (*)

The property (*) possessed by a geometric $dist^n$ has an interesting interpretation.

$P(X > m+n \mid X > n)$ = Conditional prob that of no success achived till $(m+n)^{th}$ trial given that no success till n^{th} trial.

$P(X > m)$ = No success achived till m^{th} trial.

That means starting point immaterial.

This property of a prob. distn is known as the lack of memory property.

Negative Binomial Distn :

Consider independent Bernollian trials under identical conditions till r^{th} success is achived.

X denote the number of trial needed to achive r^{th} success first time.

X takes values $r, r+1, r+2, \ldots$

r^{th} s

$\uparrow$

$\ldots\ldots\ldots K$

$$P(X = k) = \binom{k-1}{r-1} q^{k-r} p^r.$$

$k = r, r+1, \ldots\ldots$

So the p.m.f of X is

$$f_X(x) = P(X=x) = \begin{cases} \binom{x-1}{r-1} q^{x-r} p^r, & x = r, r+1, \ldots \\ 0 & o/w. \end{cases}$$

We write $X \sim NB(r,p)$, $r \in \mathbb{N}$, $p \in (0,1)$

$\{NB(r,p) : r \in \mathbb{N}, p \in (0,1)\} \rightarrow$ two parameter family.

$$E(X) = \frac{r}{p}, \quad Var(X) = \frac{rq}{p^2}$$

$$M_X(t) = E(e^{tX}) = \sum_{k=r}^{\infty} e^{tk} \binom{k-1}{r-1} q^{k-r} p^r$$

$$= \sum_{k=r}^{\infty} \binom{k-1}{r-1} (pe^t)^{k-r} (pe^t)^r = \frac{(pe^t)^r}{(1-qe^t)^r}$$

$0 < qe^t < 1.$

Poisson Distribution :

Suppose that event E, say a phone call received at a telephone exchange is occuring randomly over a time period.

Accident occuring at a crossing over a time period.

so we consider observations / occurences / happenning observed over time / area / space.

we consider this occurence under a Poisson process provided the following assumptions are satisfied.

(I) The number of outcomes / occurences during disjoint time intervals are independent.

(II) The prob. of a single occurence during a small time interval is proportional to the length of the interval.

(iii) The prob. of more than one occurence during a small time interval is negligible.

$X(t)$ denote the no of occurence in an interval of length t.

Note. λ is the rate of arrival

Then $X(t) \Rightarrow 0, 1, 2, \ldots$

$$P(X(t) = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad k = 0, 1, 2, \ldots$$

If we consider X denote the number of occurence in a unit length interval then

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Def^n: A discrete r.v. X is said to follow a Poisson dist^n with parameter $\lambda > 0$ written as $X \sim \mathcal{P}(\lambda)$ if its support is $\{0, 1, 2, \dots\}$ and its p.m.f.

$$f_X(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & o/w. \end{cases}$$

The family $\{\mathcal{P}(\lambda) : \lambda > 0\}$ is a one parameter family.

Example: Suppose customers arrive in a shopping mall according to Poisson dist^n with rate 5 per minute (1) What is the prob that no customer come in a 1 min period.

(ii) 2 customer in a 3 min period.

Soln: (1) $P(X(1) = 0) = e^{-5}$, $\lambda = 5$

(ii) $P(X(3) = 2) = \dfrac{e^{-(5\times3)}(5\times3)^2}{2!} = \dfrac{e^{-15}(15)^2}{2!}$

We know that

$$E(X) = Var(X) = \lambda.$$

$$M_X(t) = e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}.$$

Find $E\left(X(X-1)\cdots(X-r+1)\right)$ (H.W.).

<u>Theorem</u>: Let $X \sim Bin(n, p)$. Let $n \to \infty$, $p \to 0$ $\ni$ $np = \lambda$. Then $f_X(x) \to \dfrac{e^{-\lambda}\lambda^x}{x!}$

<u>Proof</u>: $f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$

$$\cong \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

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

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

$$\to 1 \cdot \frac{\lambda^x}{x!} e^{-\lambda} \cdot 1 = \frac{e^{-\lambda}\lambda^x}{x!}.$$

EX: Consider a person who plays a series of 2500 games independently. If the prob of person winning any game is 0.002, find the prob that the person will win at least two games.

Solu^n: X denote the number of success in $n = 2500$ Bernoullis trial. $p = 0.002$. $X \sim Bin(2500, 0.002)$

Since $n = 2500$ is large and $np = 5$ is a moderate

$\lambda = np = 5$.

$Y \sim P(5)$

$$P(X \geq 2) \simeq P(Y \geq 2) = 1 - P(Y=0) - P(Y=1)$$

$$= 1 - e^{-\lambda} - \lambda e^{-\lambda} = 1 - 6 \times e^{-5}$$

$$= 0.9596.$$