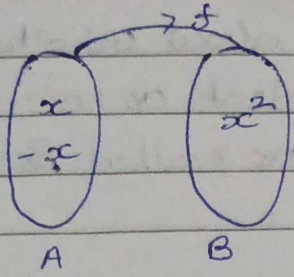
Chapter - 3

3.1 Random variable



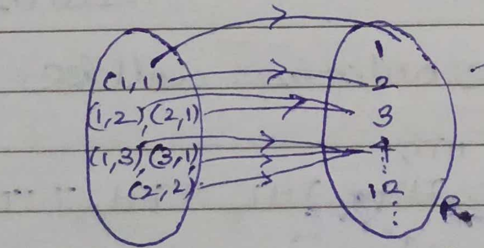
set of real no.s.

The random variable $X: S \rightarrow R$ is a function whose domain is the sample space 'S' and the range is the set of real roots.

In other words, a random variable is a rule by which we assign the real no.s to each ~~outcome~~ sample point of 'S'

Eg: In rolling two fair dice, if we observe the sum of the two faces that shows up the ~~sample sp~~ possible outcomes are $S = \{(1,1), \ldots\ldots (6,6)\}$

Let, x = the sum of the two faces.
$= 2, 3, \ldots\ldots, 12.$



It is a function.

$S = \{$Success, Failure$\}$

X(Success) $= 1$

X(Failure) $= 0$

Eg:- When a student calls a university help desk for technical support, he/she will be able to speak either immediately to someone (s for success) or will be placed on hold (f for failure), then the sample space $S = \{$Success, Failure$\}$
$= \{s, f\}$

Then defining the random variable X by $X(s) = 1$ and $X(f) = 0$

The random variable X speaks about whether the Student can speak to the help desk or not. Such type of random variables are called Bernoulli random variable.

A random variable (rv) is categorised into two types.

a) Discrete random variable (drv)

b) continuous random variable (crv)
can be rational, decimal, fractional

Discrete random variable (drv) → takes the values which are integers.

It is a random variable which takes finite no. of values that can be listed. In other words, a drv is a rv whose values constitute a finite set or a set which is countably infinite.

Eg:- Consider a statistical experiment where three fair coins are tossed once. (if a fair coin is tossed thrice.)

If the no. of heads in the outcomes will be observed, then

S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

If X = the no. of heads. → Discrete random variable

= 0,1,2,3

$2<x<3$
↓
Infinite no. of ~~variables~~ values

P(X=0) → means we are calculating the probability of getting no heads.

$= \frac{1}{8}$

P(atleast one head) $= \frac{7}{8}$

P(only one head) $= \frac{3}{8}$

Probability distribution

It tells us how the probability '1' can be distributed among all possible values of X

eg. In rolling a pair of fair dice, let X = sum of the two faces that shows up

$S = \{(1,1), (1,2), \ldots, (6,6)\} \rightarrow 36$ total outcomes

$X = 2, 3, \ldots, 12$.

$P(X=0) = 0$

$P(X=1) = 0$

$P(X=2) = 1/36$

$P(X=3) = 2/36 = \frac{1}{18}$

$P(X=4) = 3/36 = \frac{1}{12}$

(2,2), (3,1), (1,3)

$P(X=5) = 4/36 = \frac{1}{9}$

(1,4), (4,1), (2,3), (3,2)

$P(X=6) = 5/36$

(1,5), (5,1), (3,3), (2,4), (4,2)

$P(X=7) = 6/36 = \frac{1}{9}$

$P(X=8) = 5/36$

(2,6), (6,2)

$P(X=9) = 4/36 = \frac{1}{9}$

$P(X=10) = 3/36 = \frac{1}{12}$

$P(X=11) = 2/36 = \frac{1}{18}$

$P(X=12) = 1/36$

The sum of probabilities is distributed among all those probabilities. This is known as the probability distribution.

Total No. of outcomes = (No. of possibilities)$^{\text{No. of things}}$

Probability mass function (pmf) or
Probability distribution function (pdf)

For every value 'x' of a drv 'x', the pmf specifies the probability of observing that value (x) when an experiment is conducted. It may be noted that:-

i) $p(x) \geq 0$ → probability mass function denoted by small p.

ii) $\Sigma p(x) = 1$ for all possible values of x;

are the required conditions for a pmf. Here $p(x) = P(X=x)$

Q) eg: Let, X be the drv which represents the no. of heads in tossing 4 fair coins.

Calculate the pmf.

→ $2^4 = 16$ outcomes.

$X = 4, 3, 2, 1, 0$

S = {HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT}

Sum of the probabilities is equal to 1.

$p(0) = P(X=0) = 1/16$

$p(1) = P(X=1) = 4/16 =$

$p(2) = 6/16$

$p(3) = 4/16$

$p(4) = 1/16$

Cumulative distribution function (cdf)

The cdf of drv x with pmf $p(x)$ (>0) defined for each value x of X is given by

$$F(x) = P(X \leq x) = \sum_{y:\, y \leq x} p(y)$$

$$F(2) = P(X \leq 2)$$
$$= p(0) + p(1) + p(2)$$
$$= P(X=0) + P(X=1) + P(X=2)$$

Let us consider the same example as in pmf.

eg: $F(0) = P(X=0) = p(0) = 1/16$

$$F(1) = P(X \leq 1) = p(0) + p(1) = \frac{1}{16} + \frac{4}{16} = \frac{5}{16}$$
$$F(2) = P(X \leq 2) = p(0) + p(1) + p(2)$$
$$= \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16}$$

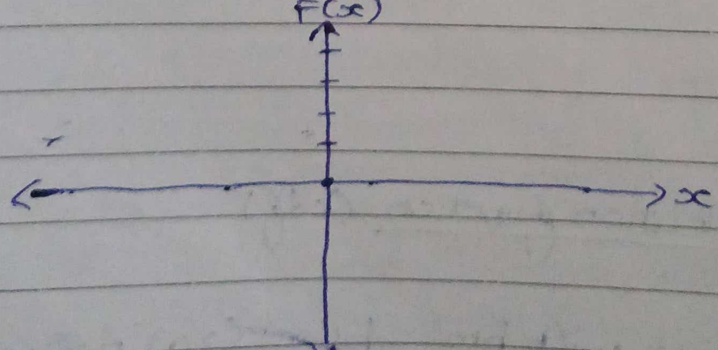
$$F(3) = p(0) + p(1) + p(2) + p(3)$$
$$= \frac{11}{16} + \frac{4}{16} = \frac{15}{16}$$

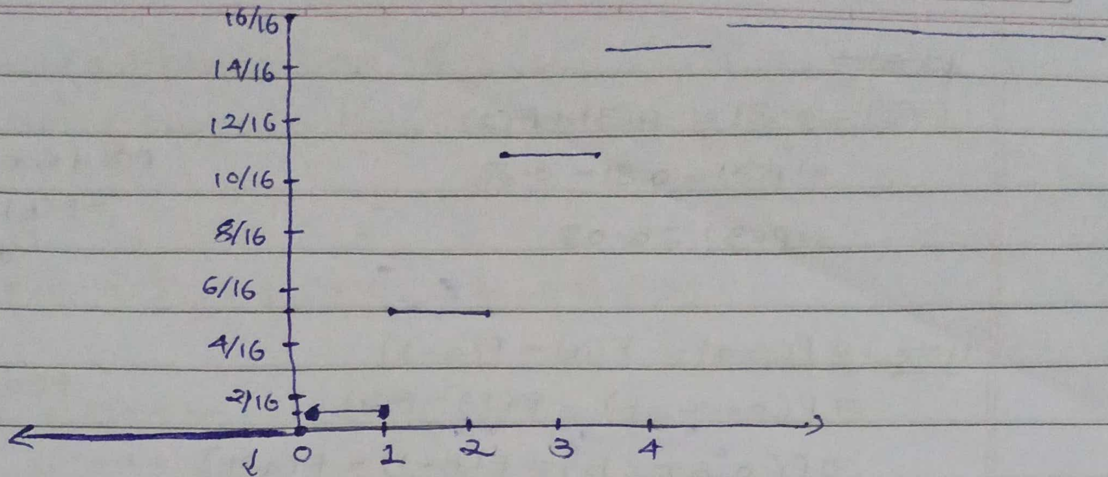
$$F(4) = 1 = p(0) + p(1) + p(2) + p(3) + p(4)$$

Define $F(x)$ for this problem.

we need to define for $-\infty$ to $+\infty$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1/16, & \text{if } 0 \leq x < 1 \\ 5/16, & \text{if } 1 \leq x < 2 \\ 11/16, & \text{if } 2 \leq x < 3 \\ 15/16, & \text{if } 3 \leq x < 4 \\ 1, & \text{if } x \geq 4 \end{cases}$$





Not touch
the origin

Note

If a and b are any two numbers with $a \le b$, then $P(a \le x \le b) = F(b) - F(a^-)$ where a^- represents a number strictly less than 'a'

If a and b are any two integers with $a \le b$, then $P(a \le x \le b) = F(b) - F(a-1)$

Q) Let, x be the no. of days of sick leave taken by a randomly selected employee of a company during a particular year of the maximum no. of allowable sick days per year is 14 and for $x = 0, 1, \ldots, 14$, $F(0) = 0.58$, $F(1) = 0.72$, $F(2) = 0.76$, $F(3) = 0.81$, $F(4) = 0.88$, $F(5) = 0.94$

$2 \le x \le 3$
$F(3) - F(1)$

Find $P(2 \le x \le 5)$ and $P(x=3) = ?$

$F(0) = P(x=0) = p(0) = 0.58$

$F(1) = 0.72 = P(x \le 1) = p(0) + p(1)$

$\Rightarrow p(1) = 0.72 - 0.58 = 0.14$

$F(2) = 0.76 = p(2) + 0.14 + 0.58 \Rightarrow p(2) = 0.04$

$F(2 \le x \le 5) = F(5) - F(2-1) = F(5) - F(1)$

$= 0.94 - 0.78$

$= 0.22$

$P(x=3) = p(3) =$

PAGE No.

DATE / /

$F(3) = 0.81 = p(3) + F(2)$

$\Rightarrow p(3) = 0.81 - 0.76$

$\Rightarrow p(3) = 0.05$

$P(a+1 \leq x \leq b)$
$= F(b) - F(a)$ if a and b are integers.

Note:- 1) $P(X=a) = F(a) - F(a-1)$

2) $P(a < x \leq b) = F(b) - F(a)$

3) $P(a \leq x < b) = F(b-1) - F(a-1)$

4) $P(a < x < b) = F(b-1) - F(a)$

$P(a \leq x \leq b-1)$

3.3 Expected value (OR Mean value, or mean or expectation)

Let, X be a discrete random variable (drv) with its possi values in $D=\{x_1, x_2, \ldots, x_n\}$ and pmf $p(x_i)$. Then the expected value of the discrete random variable X is denoted as $E(X)$ or μ_x or μ and is defined as

$$E(X) = \sum_{i=1}^{n} x_i \, p(x_i)$$

PAGE No.

DATE / /

$$= \sum_{x \in D} x \cdot p(x)$$

Note:-

i) Expectation of 1 is always equal to 1 : $E(1) = 1$

ii) $E(ax) = aE(x)$.

iii) $E(aX+b) = a \cdot E(X) + b$

iv) $E(h(X)) = \sum_{x \in D} h(x) \cdot p(x)$

where $h(x)$ is a function.

1 mark prove will come

Prove that, $E(1) = 1$

$$E(x) = \sum_{x \in D} x \cdot p(x)$$

if $p(1) = 1$

$\therefore E(1) = 0$

$$\Rightarrow E(1) = \sum_{x \in D} 1 \times p(x)$$

$$= \sum_{x \in D} p(x)$$

$$= 1$$

$$E(aX+b) = \sum_{x \in D} (ax+b)\, p(x) = a \sum_{x \in D} x\, p(x) + b \cdot \sum_{x \in D} p(x)$$

$$= a \cdot E(x) + b \times 1$$

$$= aE(x) + b$$

$\mu_{ax+b} = a\mu_x + b$ (Same thing)

Continuing classwork

~~iii) E{~~

Variance :-

Let, X be a drv with pmf $p(x)$ and expected value μ.
Then the variance of X is denoted as $V(x)$ or σ_x^2 or σ^2 and is defined as $V(x) = \sum_{x \in D} (x-\mu)^2 p(x) = E\{(x-\mu)^2\}$

$[\because E\{h(x)\} = \sum_{x \in D} h(x) \cdot p(x)]$

$D = \{x_1, x_2, \dots, x_n\}$

Note :

1. $V(ax+b) = a^2 V(x)$

or

Prove that $\sigma_{ax+b}^2 = a^2 \sigma_x^2$

2. $V(x) = E(x^2) - \{E(x)\}^2$ (shortcut formula for variance)

3. The +ve square root of variance is called standard deviation (S.D) and is denoted as σ_x or σ

i.e, $\sigma_x = \sqrt{\sigma_x^2} = \sqrt{V(x)}$

4. $\sigma_{ax+b} = |a| \sigma_x$

PAGE No.
DATE / /

Prove that, $V(ax+b) = a^2 V(x)$

$$V(ax+b) = \sum_{x \in D} \left((ax+b) - \mu_{(ax+b)}\right)^2 p(x)$$

$$= \sum_{x \in D} (ax + \cancel{b} - a \cdot \mu - \cancel{b})^2 p(x)$$

$E(ax+b) = a \cdot E(x) + b$

$\mu_{ax+b} = a \cdot \mu_x + b$

$$= a^2 \sum_{x \in D} (x-\mu)^2 p(x)$$

$$= a^2 V(x)$$

2. $V(x) = E(x^2) - \{E(x)\}^2$

$$V(x) = \sum_{x \in D} (x-\mu)^2 p(x)$$ μ_x

$$= \sum_{x \in D} (x^2 - 2x\mu + \mu^2) p(x)$$

$$= \sum_{x \in D} x^2 p(x) - 2\mu \sum_{x \in D} x p(x) + \mu^2 \sum_{x \in D} p(x)$$

$$= E(x^2) - 2\mu \times \mu + \mu^2 \times 1$$

$$= E(x^2) - 2\mu^2 + \mu^2$$

$$= E(x^2) - \mu^2$$

$$= E(x^2) - \{E(x)\}^2$$

prove that,

$\sigma_{ax+b} = |a| \sigma_x$

Chebyshev's inequality

For any probability distribution of a random variable x and any constant k (whose value is atleast 1).

~~$(P(|x-\mu|) \geq k\sigma)$~~

$$P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$ Here x is drv

$\mu \rightarrow$ The expected value of random variable x

$\sigma \rightarrow$ The Standard deviation value of x.

a) what is the value of the upper bound for k=2? → 0.25 k=3? k=4? k=5? k=10?

The upper bound for $k=2 = \frac{1}{k^2} = \frac{1}{4} = 0.25$

" " " " $k=3 = \frac{1}{k^2} = \frac{1}{9}$

" " " " $k=4 = \frac{1}{16}$

" " " " $k=5 = \frac{1}{25}$

" " " " $k=10 = \frac{1}{100}$

b)

x:	0	1	2	3	4	5	6
p(x):	0.10	0.15	0.20	0.25	0.20	0.06	0.04

Compute μ and σ. Then compute $P(|x-\mu| \geq k\sigma)$ for the values of k given in (a)

$\mu = 0.15 + 0.4 + 0.75 + 0.8 + 0.3 + 0.24$

$= 2.64$

$\sigma^2 = \sum (x-2.64)^2 \times p(x)$

$= (2.64)^2 \times 0.1 + (1.64)^2 \times 0.15 + (0.64)^2 \times 0.20 + (0.36)^2 \times 0.25$
$+ (1.36)^2 \times 0.2 + (2.36)^2 \times 0.06 + (3.36)^2 \times 0.04$

$= 2.3704$

$\sigma = \sqrt{2.3704} = 1.54$

Let us take $k=2$

$P(|x-2.64| \geq 2 \times 1.54)$

$= P(|x-2.64| \geq 3.08)$

$= 1 - P(|x-2.64| < 3.08)$

$= 1 - P\{-3.08 < x-2.64 < 3.08\}$

$= 1 - P\{-3.08+2.64 < x < 3.08+2.64\}$

$= 1 - P\{-0.44 < x < 5.72\}$

see the probability distribution table.

$= P(x=6) = 0.04$

$1-(0.1+0.15+0.2+0.25+0.2+0.06)$

$= 0.04$

c) Suppose x have the possible values $-1, 0, 1$ with probabilities $\frac{1}{18}, \frac{8}{9}, \frac{1}{18}$ respectively. Find $P(|x-\mu| \geq 3\sigma)$

~~P(|x-2.64~~

x:	-1	0	1
$p(x)$:	$1/18$	$8/9$	$\frac{1}{18}$

$\mu = -1 \times \frac{1}{18} + \frac{1}{18} = 0$

$E(x^2) = 1 \times \frac{1}{18} + \frac{1}{18} = \frac{2}{18} = \frac{1}{9}$

$\sigma^2 = E(x^2) - \{E(x)\}^2 = \frac{1}{9} - 0 = \frac{1}{9}$

$\sigma = \frac{1}{3}$

$1 - P(|x-\mu| < 3\sigma)$ $P(-1 \geq x \geq 1)$

$= 1 - P(|x-1| < 1)$ $= \frac{1}{18} + \frac{1}{18} = \frac{1}{9}$

$= 1 - P(-1 < x < 1)$

$= 1 - \frac{1}{18}$ $1 - \frac{16}{18}$

$= \frac{12}{18}$ $= \frac{1}{9}$

Moment about the origin (k th central)

The kth moment about the origin is donated as $E(x^k)$ and is defined as

$$E(x^k) = \sum_{x \in D} x^k p(x)$$

, x is the dov pmf is the probab mass function

$= E(x)$ when $k=1$

Moment about the mean

The kth moment about the mean 'μ' is given by $E[(x-\mu)^k]$

$$= \sum_{x \in D} (x-\mu)^k p(x)$$

$k=2$, it will give us $V(x)$

Moment generating function → It gives us the moment.

For a random variable 'x', then the moment generating function is donoted as $M_x(t)$ and is defined as

$$M_x(t) = E(e^{tx}) = \sum_{x \in D} e^{tx} p(x)$$

Note:- $E(x^k) = \left\{ \frac{\partial^k}{\partial t^k} M_x(t) \right\}_{t=0}$ kth partial derivative. $\frac{\partial^1}{\partial t}$

$$\left\{ \frac{\partial}{\partial t} M_x(t) \right\}_{t=0} = \left\{ \sum_{x \in D} \frac{\partial}{\partial t} (e^{tx} p(x))_{t=0} \right\}$$

$$= \left\{ \sum_{x \in D} x e^{tx} p(x) \right\}_{t=0}$$

$$= \sum_{x \in D} x p(x) = E(x)$$ when $k=1$

$$\left\{\frac{\partial^2}{\partial t^2} M_x(t)\right\}_{t=0} = \left\{\sum_{x\in D} x^2 e^{tx} p(x)\right\}_{t=0}$$

$$= \sum_{x\in D} x^2 p(x) = E(x^2) \quad k=2$$

3.4 Binomial distribution (It is corresponding to a drv)

Binomial experiment

An experiment is said to be a binomial experiment if it satisfies the following conditions.

1. The experiment consists of 'n' trials where n is fixed in advance of the experiment.
2. All trials in the experiment are independent.
3. Each outcome of the experiment results in one of the possible outcomes: success or failure.

If $p \rightarrow$ represents the probability of getting success and $q \rightarrow$ represents the probability of getting failure.

then $p+q=1$

$\Rightarrow q = 1-p$.

Binomial distribution:

derivation imp.

Let, 'A' be an event that occurs in x no. of trials where n is the total no. of trials in the experiment.

For $X \sim B(n;p)$, the pmf is given by $p(x) = {}^nC_x \times p^x \times q^{n-x}$

$= b(x;n,p)$

$p \rightarrow$

Random variable which follows the binomial distribution

$x = 0,1,2,\ldots,n$

The moment generating function for a random variable X is: $M_x(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx}\, b(x;n,p)$ → pmf

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

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

$$(p+q)^n = p^n + q^n + p^{n-1}q + p^{n-2}q^2 + \dots$$

$$=$$

$$(a+b)^n = a^n + {}^nc_1 a^{n-1} b^1 + {}^nc_2 a^{n-2} b^2 + \dots + {}^nc_r a^{n-r} b^r + \dots + b^n$$

$r = x$

$$= \sum_{r=0}^{n} {}^nc_r (a)^r (b)^{n-r}$$

$$M_x(t) = E(e^{tx}) \quad = (pe^t + q)^n$$

$$\text{Now, } \left\{\frac{\partial}{\partial t} M_x(t)\right\}_{t=0} = \left\{\frac{\partial}{\partial t}(pe^t+q)^n\right\}_{t=0}$$

$$= \left\{ \left\{ n(pe^t+q)^{n-1} \right\} \times pe^t \right\}_{t=0}$$

$$= n(p+q)^{n-1} \times p \quad \text{(Putting } t=0\text{)}$$

$$= np \quad (\text{As } p+q=1)$$

$$\therefore E(x) = \left\{\frac{\partial}{\partial x}(M_x(t))\right\}_{t=0} = np$$

$$V(x) = E(x^2) - \{E(x)\}^2$$

$$\text{Now, } E(x^2) = \left\{\frac{\partial^2}{\partial t^2} M_x(t)\right\}_{t=0} = \left[\frac{\partial}{\partial t}\left\{n(pe^t+q)^{n-1} pe^t\right\}\right]_{t=0}$$

$$= \left[np \frac{\partial}{\partial t}\left\{(pe^t+q)^{n-1} e^t\right\}\right]_{t=0}$$

$$= np\left[(n-1)(pe^t+q)^{n-2} \times pe^t + (pe^t+q)^{n-1} \times e^t\right]_{t=0}$$

when $t=0$

$$= np[(n-1)(p+q)^{n-2} \times p + (p+q)^{n-1}]$$

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

$= np[np - p + 1]$

$= np[np+q]$ $\quad 1-p=q$

$= n^2p^2 + npq$

$\therefore V(x) = E(x^2) - \{E(X)\}^2$

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

$= npq$

S.D. $(\sigma_x$ or $\sigma) = \sqrt{V(x)} = \sqrt{npq}$

$F(x) = P(X \le x)$ cumulative distribution function.

<u>Cumulative binomial distribution (or cumulative distribution function for binomial distribution</u>

For $X \sim B(n;p)$; the cdf is denoted as $B(x;n,p)$ and it is defined as:-

$$B(x;n,p) = P(X \le x) = \sum_{y=0}^{x} b(y;n,p)$$

(Coming under dsv)

3.5) Hypergeometric and negative binomial distribution

Hypergeometric distribution

1000 spoons

DF 300   NDF 700

$P(X=x)$

$= \frac{300_{C_x} \times 700_{C_{200-x}}}{1000_{C_{200}}}$

n = 200

no. of

N = No. of spoons

M DF $N-M$ (NDF)

For n trials

$P(X=x)$

$$= \frac{M_{C_x} \; {}^{N-M}C_{n-x}}{N_{C_n}} = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Let, n be the size of a random sample consisting of 'M' successes (S) and (N−M) failures (F)

If the random variable 'x' represents the no. of success in n trials, then the corresponding probability distribution is called hypergeometric distribution.

The probability mass function of the hypergeometric distribution is denoted as $h(x; n, M, N)$ and is defined as

$$h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} = p(x) \; ; \; x = 0, 1, 2, \ldots, n$$

The mean of the hypergeometric distribution is $n\frac{M}{N}$ and the variance ~~of~~ is given by $\left(\frac{N-n}{N-1}\right) n \frac{M}{N}\left(1-\frac{M}{N}\right)$

Q) Find the mean and the variance of the hypergeometric distribution.

Ans: Mean $= E(x) = \sum_{x=0}^{n} x\,p(x) = \sum_{x=0}^{n} x \cdot \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$

$$x\binom{M}{x} = x \times \frac{\lfloor M}{\lfloor x \times \lfloor M-x} = \frac{\lfloor M}{\lfloor x-1 \times \lfloor M-x} = \frac{M \lfloor M-1}{\lfloor x-1\ \lfloor M-x}$$
$$= M\binom{M-1}{x-1}$$

$$\binom{N}{n} = \frac{\lfloor N}{\lfloor n\ \lfloor N-n} = \frac{N \lfloor N-1}{n \lfloor n-1\ \lfloor N-n}$$
$$= \frac{N}{n}\binom{N-1}{n-1}$$

$\therefore$ Expectation of x

$$\text{Mean} = E(x) = \sum_{x=1}^{n} x \cdot \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

because for $x=0$, $E(x)=0$

$$E(x) = \sum_{x=1}^{n} \frac{M\binom{M-1}{x-1}\binom{N-M}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{Mn}{N}\sum_{x=1}^{n}\frac{\binom{M-1}{x-1}\binom{N-M}{n-x}}{\binom{N-1}{n-1}}$$

$M-1=P$, $x-1=y$ $\Rightarrow y=0$, when $x=1$, $y=n-1$, when $x=n$

$N-1=Q$, $n-1=m$ $\Rightarrow y=m$, when $x=n$

$N-M=Q-P$

$n-x=m-y$

$$\therefore E(x) = \frac{Mn}{N}\sum_{y=0}^{m}\frac{\binom{P}{y}\binom{Q-P}{m-y}}{\binom{Q}{m}}$$

$$= \frac{nM}{N}\left(\text{As } \sum_{y=0}^{m}\frac{\binom{P}{y}\binom{Q-P}{m-y}}{\binom{Q}{m}} = 1\right) \quad \sum_{x=0}^{n} p(x) = 1$$

PAGE No.
DATE / /

$$E(x) = \frac{nM}{N}$$

$$V(x) = E(x^2) - \{E(x)\}^2$$

$$= E(x(x-1)+x) - \{E(x)\}^2$$

$$= E(x(x-1)) + E(x) - \{E(x)\}^2$$

$$E(x(x-1)) = \sum_{x=0}^{n} \frac{x(x-1)\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$= \sum_{x=2}^{n} \frac{x(x-1)\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

$$x(x-1)\binom{M}{x} = x(x-1) \times \frac{\lfloor M}{\lfloor x \times \lfloor M-x}$$

$$= \frac{M \times M-1 \times \lfloor M-2}{\lfloor x-2 \times \lfloor M-x}$$

$$= M(M-1)\binom{M-2}{x-2}$$

$$\binom{N}{n} = \frac{\lfloor N}{\lfloor n \times \lfloor N-n} = \frac{N(N-1)\lfloor N-2}{n(n-1)\lfloor n-2 \lfloor N-n}$$

$$= \frac{N(N-1)}{n(n-1)} \times \binom{N-2}{n-2}$$

$$E(x(x-1)) = \sum_{x=2}^{n} \frac{M(M-1)\binom{M-2}{x-2} \times \binom{N-M}{n-x}}{\frac{N(N-1)}{n(n-1)} \times \binom{N-2}{n-2}}$$

$$= \sum_{x=2}^{n} \frac{nM(M-1)(n-1)}{N(N-1)} \frac{\binom{M-2}{x-2}\binom{N-M}{n-x}}{\binom{N-2}{n-2}}$$

$M-2 = R$ $\quad x-2 = Z$

$N-2 = S$ $\quad n-2 = V$

$N-M = S-R$

$n-x = V-Z$

$Z = V$, when $x = n$

$Z = 0$, when $x = 2$

$$E(x(x-1)) = \sum_{z=0}^{v} \frac{nM(M-1)(n-1)}{N(N-1)} \times \frac{\binom{R}{z}\binom{S-R}{v-z}}{\binom{S}{v}}$$

$$= \frac{nM(M-1)(n-1)}{N(N-1)} \left(\text{As } \sum_{z=0}^{v} \frac{\binom{R}{z}\binom{S-R}{v-z}}{\binom{S}{v}} = 1 \right) \text{ as } \Sigma p(x) = 1$$

$$\therefore V(x) = \frac{nM(M-1)(n-1)}{N(N-1)} + \frac{nM}{N} - \frac{n^2M^2}{N^2}$$

$$= \frac{nM}{N}\left(\frac{(M-1)(n-1)}{(N-1)} + 1 - \frac{nM}{N} \right)$$

$$= \frac{nM}{N}\left(\frac{Mn - M - n + 1}{N-1} + 1 - \frac{nM}{N} \right)$$

$$= \frac{nM}{N}\left(\frac{M\cancel{N}n - MN - nN + \cancel{N} + N^2 - \cancel{N} - nM\cancel{N} + nM}{N(N-1)} \right)$$

$$\cancel{= \frac{nM}{N}\left(nM - MN - nN \right.}$$

$$= \frac{nM}{N}\left(\frac{N^2 - MN - nN + nM}{N(N-1)} \right)$$

$N^2 - Nn - MN + Mn$

$$= \frac{nM}{N^2(N-1)} \left((N-M)(N-n) \right)$$

$$= \frac{nM}{N(N-1)} \left(1 - \frac{M}{N}\right)(N-n)$$

$$= \left(\frac{N-n}{N-1}\right) n \frac{M}{N} \left(1 - \frac{M}{N}\right)$$

If $p = \frac{M}{N}$,

$$\therefore E(x) = np$$

$$V(x) = \left(\frac{N-n}{N-1}\right) np(1-p)$$ → Difference in variance

Negative binomial distribution

The -ve binomial rv and its distribution based on an experiment are satisfying the following conditions.

a. The experiment consist of a sequence of a sequence of a sequence of independent trials.

b. Each trial results in one of the two possible outcomes (Success (S), Failure (F)).

c. The probability of success is constant from trial to trial The trial i.e; $P(S \text{ on trial } i) = p$; $i = 1, 2, 3, \ldots$

d. The experiment continues (or the no. of trials to be performed continues) till 'r' no. of 'successes' have been observed where 'r' is a specified +ve integer

The pmf of the -ve Binomial rv 'X' with 'r' successes and $p(S)=p$, is given by $nb(x;r,p) = \binom{x+r-1}{r-1} p^r (1-p)^x$; $x=0,1,\dots$

$[{}^nC_r = {}^nC_{n-r}]$

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

The mean and the variance of -ve binomial distribution are $E(x) = \frac{r(1-p)}{p}$ & $V(x) = \frac{r(1-p)}{p^2}$

Prove this

Derivation for E(x)

we know

$$E(x) = \sum x \cdot p(x) = \sum_{x=0}^{\infty} x \cdot nb(x;r,p)$$

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

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

$$= \sum_{x=1}^{\infty} x \times \frac{\lfloor x+r-1}{\lfloor x \times \lfloor r-1} \times p^r \times (1-p)^x$$

$$= \sum_{x=1}^{\infty} \frac{r \times \lfloor x+r-1}{\lfloor x-1 \ \lfloor r} p^r \times (1-p)^x$$

~~$x+r-x = x+x$~~

$x+r-1-x+1$

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

Let, $x-1=y$

$r = s-1$

$$= r \sum_{y=0}^{\infty} \binom{s+y-1}{y} (p)^{s-1} \times (1-p)^{y+1}$$

$$= \frac{r(1-p)}{p} \sum_{y=0}^{\infty} \binom{s+y-1}{y} p^s (1-p)^y$$

PAGE No.
DATE / /

$\sum p(x) = 1$

$\therefore E(x) = \frac{r(1-p)}{p}$

$E(x(x-1))$

Prove that $V(x) = \frac{r(1-p)}{p^2}$ (HW)

- $E(x) = \frac{r(1-p)}{p}$

* The expected no of trials until we get the 1st success is $\frac{1}{p}$

The expected no of failures until we get the 1st success
$= \frac{1}{p} - 1 = \frac{1-p}{p}$

The expected no of failures until we observe r success
$r\left(\frac{1-p}{p}\right)$

Ex: 3.4

PAGE No.
DATE / /

Continuing classwork

Proof for $V(x) = \frac{r(1-p)}{p^2}$

we know that, $V(x) = E(x^2) - \{E(x)\}^2$

$= E(x(x-1) + x) - \{E(x)\}^2$

$= E(x(x-1)) + E(x) - \{E(x)\}^2$

$E(x(x-1)) = \sum x(x-1)\, p(x(x-1)) = \sum_{x=0}^{\infty} x^{(x-1)}\, nb(x;\, r, p)$

we know that, $nb(x; r, p) = \binom{x+r-1}{x} p^r \times (1-p)^x$

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

$\frac{\lfloor x+r-1}{\lfloor x \times \lfloor x+r-1-x}$

$= \sum_{x=0}^{\infty} \cancel{x}(\cancel{x}-1) \times \frac{\lfloor x+r-1}{\cancel{x}(\cancel{x}-1)\lfloor x-2 \times \lfloor r-1} \times p^r \times (1-p)^x$

$\frac{\lfloor x+r-1}{\lfloor x \times \lfloor r-1}$

$= r(r+1) \sum_{x=2}^{\infty} \binom{x+r-1}{x-2} p^r (1-p)^x$

$\cancel{x}+r-1-\cancel{x}+2$

$r+1$

$x+r-1$
$(x-2)+r-1$

$x-2 = y$
$r = S-2$

$= r(r+1) \sum_{y=0}^{\infty} \binom{S+y-1}{y} p^{(S-2)} (1-p)^{y+2}$

8 (Block chain and cybersecurity) x

$x+r-1$
$(y+\cancel{2})+(S-\cancel{2})-1$
$y+S-1$

$= \frac{r(r+1)(1-p)^2}{p^2} \sum_{y=0}^{\infty} \binom{S+y-1}{y} p^S (1-p)^y$

Q)

$y = x$
$S = r$

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

$S-2+1$
$(S-1)$

$\therefore V(x) = E(x(x-1)) + E(x) - \{E(x)\}^2$

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

$= \frac{r(1-p)}{p}\left(\frac{(r+1)(1-p)}{p} + 1 - \frac{r(1-p)}{p}\right)$

$= \frac{r(1-p)}{p}\left(\frac{(r+1)(1-p) + (r+1)(1-p) - r(1-p)}{p}\right)$

$$= \frac{r(1-p)}{p}\left(\frac{r - rp + 1 - p + r - rp + 1 - p - r + rp}{p}\right)$$

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

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

$$\Rightarrow V(x) = \frac{r(1-p)}{p^2}$$ Hence, proved.

Note

When $r=1$, the pmf of -ve binomial distribution becomes

$$nb(x; r, p) = \binom{x+r-1}{r-1} p^r (1-p)^x$$

$$nb(x; 1, p) = p(1-p)^x \quad [\because n_{c_0} = 1]$$

which represents the probability mass function of the geometric distribution.

$$\text{Mean} = E(x) = \frac{r(1-p)}{p} = \frac{(1-p)}{p} \quad [\because r=1]$$

$$\text{Variance} = V(x) = \frac{r(1-p)}{p^2} = \frac{1-p}{p^2} \quad [\because r=1]$$

No. of

Ex: 3.6

Poisson's distribution →(when n is very very large)

A drv 'x' is said to have poisson's distribution if with parameter 'μ'(>0) if the pmf of X is given by $p(x;\mu)$

$$= \frac{e^{-\mu}\mu^{x}}{\lfloor x}\ ;\ x = 0, 1, 2, \ldots\ldots$$

The mean and variance of poisson's distribution are given by

$E(x) = V(x) = \mu = np$ — n is very large but represents the no. of trials.
p → no. of successes.

Derivation of mean and variance of poisson's distribution

$$M_x(t) = E(e^{tX}) = \sum_{x=0}^{\infty} \left(e^{xt} \frac{e^{-\mu} \mu^x}{\lfloor x} \right)$$

$$= \sum_{x=0}^{\infty} e^{-\mu} \frac{(\mu e^t)^x}{\lfloor x}$$

$$= e^{-\mu} \left\{ 1 + \mu e^t + \frac{(\mu e^t)^2}{\lfloor 2} + \frac{(\mu e^t)^3}{\lfloor 3} + \dots \right\}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

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

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

Derive the same for mean and variance.

Ex: 3.5

PAGE No.	
DATE	/ /

$$E(x) = \frac{\partial}{\partial t}\left(M_x(t)\right) = \left[\frac{\partial}{\partial t}\left(e^{\mu(e^t - 1)}\right)\right]_{t=0}$$

$$\left[\frac{\partial}{\partial t}\left(e^{\mu e^t} / e^{\mu}\right)\right]_{t=0}$$

$$= \left[\frac{1}{e^{\mu}} \times e^{\mu e^t} \times \mu e^t\right]_{t=0}$$

when, $t=0$

$$E(x) = \frac{1}{e^{\mu}} \times e^{\mu} \times \mu = \mu$$

$$E(x^2) = \left\{\frac{\partial^2}{\partial t^2}\left(M_x(t)\right)\right\}_{t=0}$$

$$= \left\{\frac{\partial}{\partial t}\left(\frac{\partial}{\partial t}(M_x(t))\right)\right\}_{t=0}$$

$$\frac{\partial}{\partial t}\left\{\left(\frac{1}{e^{\mu}} \times e^{\mu e^{t}} \times \mu e^{t}\right)\right\}_{t=0}$$

$$= \frac{\mu}{e^{\mu}} \times \left\{\frac{\partial}{\partial t}\left(e^{\mu e^{t}} \times e^{t}\right)\right\}_{t=0}$$

$$= \frac{\mu}{e^{\mu}} \times \left[e^{\mu e^{t}} \times e^{t} + e^{t} \times e^{\mu e^{t}} \times \mu e^{t}\right]_{t=0}$$

$$= \frac{\mu}{e^{\mu}} \times \left\{e^{\mu} + e^{\mu} \cdot \mu\right\}$$

$$= \mu(1+\mu)$$

$$= \mu + \mu^2$$

$$V(x) = E(x^2) - \{E(x)\}^2$$

$$= \mu + \mu^2 - \mu^2$$

$$= \mu$$

$$= np$$

$\therefore E(x) = V(x) = \mu$

Note:- when 'n' is large $n \to \infty$

$b(x; n, p) \to p(x; \mu)$ (The binomial distribution tends to poisson's distribution)

(Prove this)

PAGE No.
DATE / /

Note 2) The no. of events occuring at a particular time interval (t) is a poisson's process
The expected no. of events in that interval of time (in time 't') is given by

$$E(x) = \mu = \alpha t \quad , \text{ where '}\alpha\text{' specifies the rate.}$$

Ex: 3.6

$$F(x;\mu) = \sum_{y=0}^{x} \frac{e^{-\mu}\mu^{y}}{y!}$$

$\mu = 1 - F(x)$

Continuing classwork

IMP $b(x;n,p) \rightarrow p(x;\mu)$ when $n \rightarrow \infty$

Q) Show that poisson's distribution is the limiting case of binomial distribution.

i.e, $\lim_{n \rightarrow \infty} b(x;n,p) = p(x,\mu)$

Pf. $b(x;n,p) = {}^nC_x \, p^x (q)^{n-x}$

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

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

$\lfloor n$ = Product of n consecutive natural numbers.

$= \frac{(n \times (n-1) \times (n-2) \times \ldots (n-x+1)) \times \cancel{\lfloor n-x}}{\lfloor x \times \cancel{\lfloor n-x}} p^x (1-p)^{n-x}$

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

$\mu = np$

$\Rightarrow n = \frac{\mu}{p}$

$= \frac{\left(\frac{\mu}{p}\right)^x \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) \cdots \left(1-\frac{x+1}{n}\right) \cancel{p^x} (1-p)}{\cancel{p^x} \lfloor x}$

$$\frac{\mu^x\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x+1}{n}\right)(1-p)^{n}}{\lfloor x \qquad (1-p)^x}$$

$$= \frac{\mu^x\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x+1}{n}\right)}{\lfloor x} \quad \frac{\left(1-\frac{\mu}{n}\right)^n}{(1-p)^x}$$

when $n \to \infty$ i.e, when $p \to 0$

$$\lim_{n\to\infty} b(x;n,p) = \frac{\mu^x}{\lfloor x} \quad \frac{\lim\limits_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n}{\lim\limits_{p\to 0}(1-p)^x}$$

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$$

$$= \frac{\mu^x}{\lfloor x} \times \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n$$

$$= \frac{\mu^x}{\lfloor x} \times \lim_{n\to\infty}\left(\left(1-\frac{\mu}{n}\right)^{-\frac{n}{\mu}}\right)^{-\mu}$$

$$= \frac{\mu^x}{\lfloor x}\, e^{-\mu} = p(x;\mu)$$