DS & AI
CS & IT

Statistics - I

(Discrete Random Variable)

Lecture No. 03

By- Dr. Puneet Sharma Sir

Recap of previous lecture

Topic

PROBABILITY DISTRIBUTION

Topics to be Covered

Topic

(1) Geometric Distribution

(2) Binomial Distribution

Thumb Rule of this Chapter → Try to avoid making Question by using following words;

"If, what if, AGAR, YADI, TOH, ..."

OR

Don't Try to develop Question (by your little mind until you have a complete understanding of the chapter) & try to solve the Quest.

Random Variable

RECAP

D.R.V (X)

↓

Discrete Prob Distribution

eg (Geometric, Binomial, Poisson)

↓

Prob Mass func^n (p.m.f) $= p_i$

$p_i \geq 0$, $\sum p_i = 1$

C.R.V (x)

↓

Continuous Prob Distribution

eg (Exponential, uniform, Normal)

↓

Prob. Density func^n (p.d.f) $= f(x)$

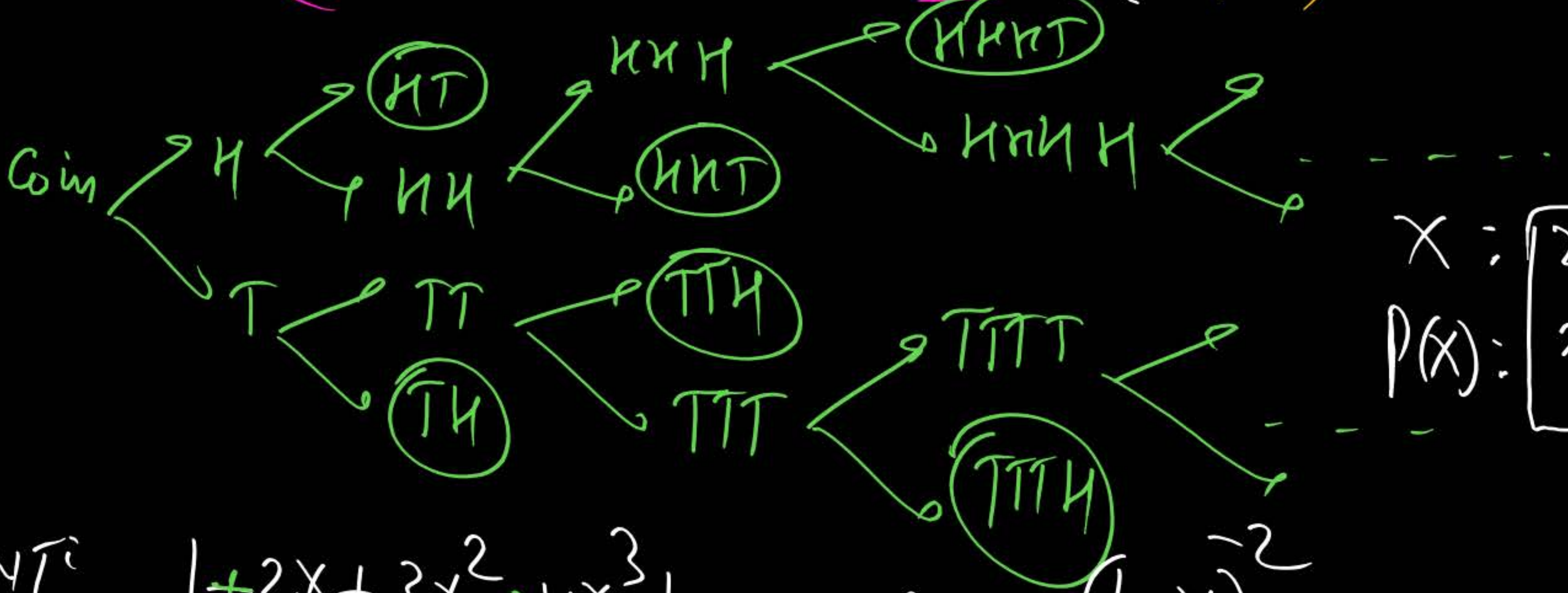
$f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$

A fair coin is tossed repeatedly till both head and tail appear at least once. The average number of tosses required is______.

Ans = 3

Min tosses Req = 2
Max " " = No idea
Av " " = ??

$X = \{\text{No. of tosses Required}\} = \{2, 3, 4, \ldots\ldots\}$



$X:$	2	3	4 ......
$P(X):$	?	?	?

HINT: $1 + 2x + 3x^2 + 4x^3 + \ldots\ldots \infty = (1-x)^{-2}$

$p_1 = P(X = 2 \text{ tosses}) = P(HT \text{ or } TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$p_2 = P(X = 3 \text{ tosses}) = P(HHT \text{ or } TTH) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$

$p_3 = P(X = 4 \text{ tosses}) = P(HHHT \text{ or } TTTH) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$

& so on ......

X :	2	3	4	5	......
$P(x)$:	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	......

$\because \sum p_i = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots\ldots$

$= \frac{1/2}{1 - 1/2} = 1$

$S = \{ \underbrace{HT, TH}_{X=2}, \underbrace{HHT, TTH}_{X=3}, \underbrace{HHHT, TTTH}_{X=4}, \ldots\ldots \}$

$E(X) = \sum p_i x_i = p_1 x_1 + p_2 x_2 + p_3 x_3 + \ldots\ldots$

$= \frac{1}{2}(2) + \frac{1}{4}(3) + \frac{1}{8}(4) + \ldots\ldots$

$$E(X) = 2\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + 4\left(\frac{1}{2}\right)^3 + \cdots\cdots$$

$$= \left[1 + 2\left(\frac{1}{2}\right) + 3\left(\frac{1}{2}\right)^2 + 4\left(\frac{1}{2}\right)^3 + \cdots\cdots\right] - 1$$

$$\left[\because\ 1 + 2x + 3x^2 + 4x^3 + \cdots\cdots = (1-x)^{-2}\right]$$

$$= \left[\left(1-\frac{1}{2}\right)^{-2}\right] - 1 = \left(\frac{1}{2}\right)^{-2} - 1 = \frac{1}{2^{-2}} - 1 = 2^2 - 1 = 4 - 1 = 3$$

(*) $(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots\cdots$

$n = -$ve integer OR fraction

eg $(1-x)^{-2} = 1 + (-2)(-x) + \frac{(-2)(-2-1)}{2!}(-x)^2 + \frac{(-2)(-2-1)(-2-2)}{3!}(-x)^3 + \cdots\cdots$

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

eg $(1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + \cdots\cdots$

Q. Each of the 9 words in the sentence are written on a separate pieces of paper.

"The Quick Brown fox jumps over the Lazy Dog"

$\underbrace{\text{The}}_{=3}$ $\underbrace{\text{Quick}}_{=5}$ $\underbrace{\text{Brown}}_{=5}$ $\underbrace{\text{fox}}_{=3}$ $\underbrace{\text{jumps}}_{=5}$ $\overbrace{\text{over}}^{4}$ $\underbrace{\text{the}}_{=3}$ $\overbrace{\text{Lazy}}^{4}$ $\underbrace{\text{Dog}}_{=3}$

These pieces are kept in a Box. If one piece is drawn at Random then find Expected length of word drawn?

(a) 3.89

(b) 1.33

(c) 4.00

(d) 5

$X = \{\text{length of word drawn}\} = \{3, 4, 5\}$

- Min length = 3
- Max length = 5
- Av. length = ?

X:	3	4	5
$P(X)$:	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{3}{9}$

So $E(X) = \sum p_i x_i$

$= \frac{4}{9}(3) + \frac{2}{9}(4) + \frac{3}{9}(5) = \frac{35}{9} = 3.89$

Note: $\bar{X} = \frac{\sum X}{N} = \frac{3+3+3+3+4+4+5+5+5}{9} = 3.89$

Average length of word drawn = 3.89.

BERNOULLI TRIAL (Bernoullie R Exp) → whenever in a Random Exp we have only two possible outcomes then such Types of Exp are called Bernoullie Trials.

Here we will assume one outcome as success and another outcome as failure

i.e R Exp → success, failure & $p = \text{Prob(success)}$, $q = \text{Prob(failure)}$ where $\boxed{q + p = 1}$

Note (1) Geometric Dist, Binomial Dist, Poisson Dist are based on Bernoullie Trials.

(2) Success = which is Required should be assumed as success.

eg A Die is thrown then success & failure can be assumed as success $= \{6\}$, failure $= \{1, 2, 3, 4, 5\}$ i.e $p = \frac{1}{6}$, $q = \frac{5}{6}$

Geometric Distribution (−ve Binomial Dist) →

→ this Distribution helps to find Number of Trials Required before getting 1st success

→ this Dist is also Based on Bernoullie Trials.

→ $(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots$

where $n = -$ve integer or any fraction

eg $(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots$

eg $\boxed{(1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \cdots}$

eg $(1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + \cdots$

eg $(1-x)^{2} = \cancel{1 - 2x + 3x^2 - 4x^3 + \cdots}$ Blunder.

so $(1-x)^2 = 1 - 2x + x^2$

Geometric Distribution → This Distribution helps to find Number of Trials required to get 1^{st} success.

Let $X = \{$Number of Trials Required to get 1^{st} success$\}$

& $p = P(\text{getting success})$ & $q = P(\text{getting failure})$, where $q+p=1$

Prob Dist:

X :	1	2	3	4	5	.....	$(r+1)^{th}$	......∞
$P(X)$:	p	qp	q^2p	q^3p	q^4p	.....	$(q^r p)$	.....

$\because \sum p_i = p + qp + q^2p + q^3p + \ldots\ldots$

$= p(1+q+q^2+q^3+\ldots\ldots) = p\cdot\left(\frac{1}{1-q}\right) = p\left(\frac{1}{p}\right) = 1$ i.e. all correct ✓

& $P(X=r^{th}) = P(\text{getting } 1^{st} \text{ success in } r^{th} \text{ trial}) = \boxed{q^{r-1}\cdot p}$ Here $X \sim G\{p\}$

① $\text{Mean}(X) = E(X) = \sum p_i x_i = p_1x_1 + p_2x_2 + p_3x_3 + \dots \infty$

$$= p(1) + qp(2) + q^2p(3) + q^3p(4) + \dots$$

$$= p\underbrace{\left[1 + 2q + 3q^2 + 4q^3 + \dots\right]} = p\left[(1-q)^{-2}\right]$$

$\text{Mean}(X) = \dfrac{p}{(1-q)^2} = \dfrac{p}{p^2} = \dfrac{1}{p}$ Learn. is $\boxed{\text{Av of G.R.V} = \dfrac{1}{p}}$

② $\text{Var}(X) = E(X^2) - (E(X))^2 = (\text{Do yourself}) - \left(\dfrac{1}{p}\right)^2 = \dots = \dfrac{q}{p^2}$ Learn.

③ $S.D(\sigma) = +\sqrt{\text{Var}(X)} = \dfrac{\sqrt{q}}{p}$.

Q. Suppose you are playing a Game of Dart and the probability of getting success is 0.4 then find the Prob that you will hit the Bull's Eye in 4^{th} Trial ?

M-I (Using Common Sense) →

$$\text{Req Prob} = P(F) \cdot P(F) \cdot P(F) \cdot P(S)$$
$$= (0.6)^3 \times (0.4)$$
$$= 0.0864$$

M-II (Using Geometric Dist) →

$X = \{\text{No. of trials Required to get } \underbrace{\text{Bull's Eye}}_{\text{Success}}\}$

$p = P(\text{Success}) = P(\text{Bull's Eye}) = 0.4$

$q = P(\text{failure}) = P(\text{Not getting B. Eye}) = 0.6$

$$\text{Req Prob} = \boxed{P(X=4) = q^3 p} = (0.6)^3 (0.4)$$
$$= 0.0864$$

Q: Consider a Company that produces on an Average 3 defective Bulbs out of 60 Bulbs. Then Find the prob that 1^{st} def Bulb will be found when 6^{th} one is tested?

= success.

Sol: $X = \{$ Number of Trials Required to get 1^{st} def. Bulb $\}$ → success.

$$p = P(\text{Def Bulb}) = \frac{3}{60} = \frac{1}{20}$$

$$q = P(\text{Non Def}) = 1 - p = 1 - \frac{1}{20} = \frac{19}{20}$$

$$P(X=6) = P(\text{we are getting } 1^{st} \text{ def Bulb in } 6^{th} \text{ testing})$$

$$= q^5 p = \left(\frac{19}{20}\right)^5 \left(\frac{1}{20}\right) = 0.0386$$

M-II Req Prob $= (\text{N.Def})^5 (\text{Def})$

$$= \left(\frac{19}{20}\right)^5 \left(\frac{1}{20}\right)$$

$$= 0.0386$$

A fair die with faces {1, 2, 3, 4, 5, 6} is thrown repeatedly till '3' is observed for the first time. Let X denote the number of times the die is thrown. then $E(X) =$ ______ $=$ Six

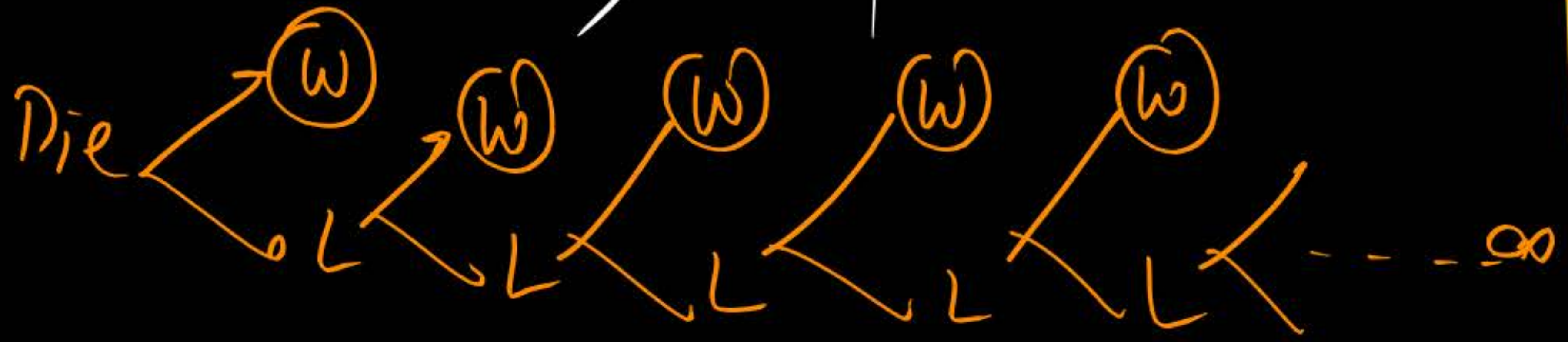
$X:$	1	2	3	4	----
$P(X)$	$\frac{1}{6}$	$\left(\frac{5}{6}\frac{1}{6}\right)$	$\left(\frac{5}{6}\right)^2\frac{1}{6}$	$\left(\frac{5}{6}\right)^3\frac{1}{6}$	------

M-I (using Common sense) $\rightarrow$

$P(W) = P(3) = \frac{1}{6}$ & $P(L) = P(\bar{3}) = \frac{5}{6}$

$S = \{W, LW, LLW, LLLW, \dots\}$

$\;\;\;\;\;\;\;\;\;\;\;\;\;1,\;\;\; 2,\;\;\;\; 3,\;\;\;\;\; 4$



$E(X) = \sum p_i x_i$

$= \frac{1}{6}(1) + \frac{5}{36}(2) + \frac{25}{216}(3) + \dots$

$= \frac{1}{6}\left[1 + 2\left(\frac{5}{6}\right) + 3\left(\frac{5}{6}\right)^2 + \dots\right]$

$= \frac{1}{6}\left[\left(1 - \frac{5}{6}\right)^{-2}\right] = \frac{1}{6}\left[\left(\frac{1}{6}\right)^{-2}\right]$

$= \frac{6^2}{6} = 6$

(M-II) Using G-Dist: $X = \{$ Number of trials Required to get 1^{st} '3' $\}$

success

$p = P(3) = \frac{1}{6}$, $q = P(\bar{3}) = \frac{5}{6}$.

we know that $\boxed{E(X) = \frac{1}{p}} = \frac{1}{(1/6)} = 6$ Ans

Q: If we are tossing $\boxed{\text{Biased Coin}}$ then to get 1st Head, Expected number of Trials are 4, then what is the prob of getting 1st Head in 2nd Trial?

(a) 1/4

(b) 3/4

(c) 1/16

(d) 3/16

$X = \{\text{Number of Trials Required to get 1st Head}\}$ success.

ATQ, $E(X) = 4$ —— (1)

But we know that, for G-Dist

$$E(X) = \frac{1}{p} \quad \text{—— (2)}$$

By (1) & (2), $\frac{1}{p} = 4 \Rightarrow p = \frac{1}{4}$

So $q = 3/4$.

$P(X=2) = P(\text{getting 1st Head in 2nd trial})$

$$= q \cdot p = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$$

General Discussion

Geometric $\leftarrow$ P(getting 1^{st} Head in 6^{th} Trial) = ? $= \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)$

Binomial.

P(" exactly 1 H in 6 Trials) = ? $\boxed{{}^6C_1 \cdot \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)} = \frac{{}^6C_1}{2^6}$

P(" " 2 H in 6 Trials) = ? ${}^6C_2 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^2$

P(" " 3 H in 6 Trials) = ? ${}^6C_3 \cdot \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3$

Explanation: → fav cases = { (HTTTTT), (THTTTT), (TTHTTT) (TTTHTT) (TTTTHT) (TTTTTH) }

$= \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^6 = 6 \times \left(\frac{1}{2}\right)^6.$

Binomial Dist

Necessary Condⁿ for B.Dist → There are four N. Conditions;

1. Number of Trials (R.Exp) should be finite i.e n = finite
2. Each Trial (R.Exp) should be Independent.
3. Each Trial (R.Exp) has only two possible outcomes known as success / failure
   i.e (Each Trial must be of Bernoullie Type)
4. The prob of success for each Trial (R Exp) should be constant.

Shortcut → whenever we are not sure about the location of success, we can apply B.Dist.

g

2B 3G	4B 1G	1B 4G	0B 5G	5B 0G	3B 2G
$p=\frac{2}{5}$	$p=\frac{4}{5}$	$p=\frac{1}{5}$	$p=0$	$p=1$	$p=\frac{3}{5}$

it is the voilation of 4^{th} N. Cond^n i.e $p \neq$ Constant.
So we can't apply B. Dist for this Question.

IMPORTANT SLIDE

FQ A coin is tossed 6 times then find

$p = P(H) = 1/2$
$q = P(T) = 1/2$

① $P(\text{getting } 1^{st} \text{ Head in } 5^{th} \text{ trial}) = ? = q^4 \cdot p \cdot (S.O) = \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)(1)$

② $P(\text{only } 1^{st} \text{ two tosses produces Head}) = ? = P(HHTTTT) = \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4$

③ $P(1^{st} \text{ two tosses produces Head}) = ?\ P(HH\ \underbrace{SO}) = \left(\frac{1}{2}\right)^2 \cdot (1)^4$

④ $P(\text{exactly 2 tosses produces Head}) = ? = {}^6C_2 \cdot p^2 q^4 = {}^6C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^4$

⑤ $P(\text{getting } 2^{nd} \text{ Head in } 6^{th} \text{ toss}) = ?\ P(\text{exactly one H in 5 tosses}) \times P(\text{H in } 6^{th} \text{ toss})$

$= \left({}^5C_1 \cdot p^1 \cdot q^4\right) \times \frac{1}{2} = \left[{}^5C_1 \times \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^4\right] \times \frac{1}{2}$

(x) A coin is tossed 10 times then find the prob of getting 4^{th} H in 9^{th} toss?

Sol: Req Prob = P(getting exactly 3H in eight tosses) × P(H in 9^{th} toss)

× P(S.O in 10^{th} toss)

$$= \left(\begin{array}{c}\text{use B. Dist.}\\ {}^{8}C_{3}\cdot p^{3}q^{5}\end{array}\right)\times\left(\frac{1}{2}\right)(1)$$

$$= \left({}^{8}C_{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{5}\right)\left(\frac{1}{2}\right)(1)$$

Def^n :- Let X is Discrete Random Variable (DRV) such that it's Probability Mass function (p.m.f) is defined as;

$$\boxed{P(X = r \text{ success}) = {}^nC_r \cdot p^r \cdot q^{n-r}}$$ where $q + p = 1$ ↗ $p = P(\text{success})$ ↘ $q = P(\text{failure})$

then X is Called Binomial Random Variable (B.R.V) & the distribution formed is Called B. Distribution with parameters n & p.

ie X can be denoted as $X \sim B\{n, p\}$

① Parameters/Statistical Attributes → those Numerical Quantities which are necessary to apply Standard Result are called Parameters

② $X = \{$which is Required should be assumed as $X\} \simeq \{$No of success$\}$

(3) Binomial Dist :

X :	0	1	2	3	4	....	$(n-1)$	n
$P(X)$:	p_0	p_1	p_2	p_3	p_4	....	p_{n-1}	p_n

where $p_i = P(X=i) = {}^nC_i\, p^i q^{n-i} \Rightarrow p_0, p_1, p_2 \cdots p_n$ can be Calculated.

$$\sum p_i = \sum_{i=0}^{n} \left({}^nC_i\, p^i q^{n-i}\right) = (q+p)^n = (1)^n = 1^{\text{finite}} = 1$$

we know that from 12th, $\boxed{\sum_{r=0}^{n} {}^nC_r \cdot x^{n-r} \cdot a^r = (x+a)^n}$

(4) Mean$(X) = E(X) = \sum p_i X_i = p_0X_0 + p_1X_1 + p_2X_2 + \cdots + p_nX_n = \text{________} = np$

(5) $Var(X) = E(X^2) - (E(X))^2 = (\text{Do yourself}) - (np)^2 = \cdots\cdots = npq$

(6) $SD(\sigma) = +\sqrt{Var(X)} = +\sqrt{npq}$

(7) In B. Dist: Mean > Variance

eg: 10 ships are going in an Atlantic ocean then find the prob that exactly 3 will come back if history suggest that out of 11000 ships only 10000 came back?

M-I (using common sense) →

$p = P(\text{ship will come back}) = \frac{10000}{11000} = \frac{10}{11}$

$q = P(\text{ship will not come back}) = \frac{1}{11}$

Req Prob $= P(\text{exactly 3 will come back})$

$= {}^{10}C_3 \cdot \left(\frac{10}{11}\right)^3 \left(\frac{1}{11}\right)^7$

M-II $n = 10$ ships (finite)

Each ship is Ind, ship → success $(=p)$, failure $(=q)$

$P(\text{success for each ship}) = \frac{10}{11} = \text{Constant}$

$n = 10,\ p = \frac{10}{11},\ q = \frac{1}{11},\ r = ? = 3.$

$X = \{\text{No. of ships coming back}\}$ → success.

$P(X = 3\ \text{success}) = {}^{10}C_3 (p)^3 (q)^7 = \text{Ans}$

Q: In a Box 10% items are Defective. If 10 items are choosen at Random then find the prob of getting exactly 2 Def items?

$X = \{\text{Number of Def. items}\} \to$ success

$p = P(\text{Def items}) = 10\% = 0.1$

$q = P(\text{Non Def items}) = 90\% = 0.9$

$n = 10$ items, $r = ? = 2$ success.

$$\boxed{P(X = r \text{ success}) = {}^{n}C_{r}\, p^{r} q^{n-r}}$$

$$P(X = 2 \text{ Def. items}) = {}^{10}C_{2}(0.1)^{2}(0.9)^{8} = 0.1937 \simeq 19\% \text{ chance.}$$

(ii) $P(\text{getting at least one Defective item}) = ? = 1 - P(\text{No Def. item}) = 1 - P(X=0) = 1 - \left({}^{10}C_{0}\, p^{0} q^{10}\right)$

$= 1 - (0.9)^{10}$

eg: If Mean & Variance of B.Dist are 4 & 12 resp then find Complete B.Dist.

Sol: Mean $= 4$, Var $= 12$ $\Rightarrow$ (Question is WRONG) $\because$ Mean $\not<$ Variance in B.Dist.

eg: If Mean & Variance of B.Dist are 12 & 4 resp then find Complete B.Dist.

Sol: Mean $= 12$, Var $= 4$

$np = 12$, $npq = 4$ & $q + p = 1$

$\Downarrow$ $\quad (12)q = 4 \quad \Downarrow$

$n\left(\frac{2}{3}\right) = 12 \quad q = \frac{1}{3} \quad p = \frac{2}{3}$

$n = 18$

So Complete B.Dist is $= (q+p)^n$

$= \left(\frac{1}{3} + \frac{2}{3}\right)^{18}$ Ans

If X is a discrete random variable that follows Binomial distribution, then which one of the following ~~response~~ relations is correct?

Recurrence

(HW)

(a) $P(r+1) = \dfrac{n-r}{r+1} P(r)$

(b) $P(r+1) = \dfrac{p}{q} P(r)$

(c) $P(r+1) = \dfrac{n+r}{r+1} \dfrac{p}{q} P(r)$

(d) $P(r+1) = \dfrac{n-r}{r+1} \dfrac{p}{q} P(r)$ ✓

$$\boxed{P(X=r+1) = \frac{n-r}{r+1} \cdot \frac{p}{q} \cdot P(X=r)}$$

NIRMAN
2025

Dr Puneet Sir pw

@DRPUNEETSIRPW

Thank

YOU