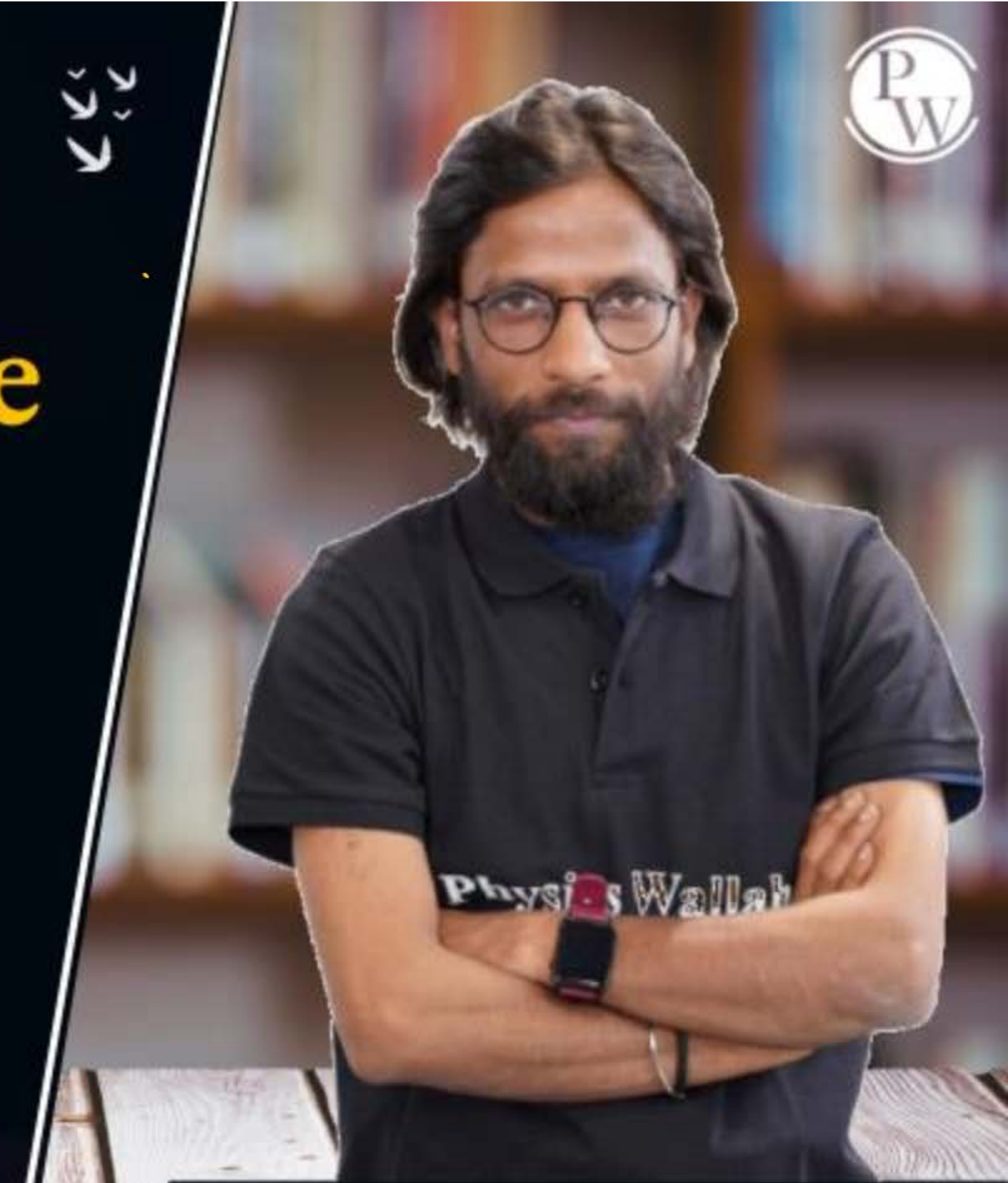
Data Science and Artificial Intelligence

Probability and Statistics

Discrete Probability Distribution

Lecture No.- 06



By- Rahul Sir

Topics to be Covered

24 — Last line

DA/DS Test/Statistics

Conditional

Prob. Distribution (Discrete)

Topic Question Based on Discrete Random Variable -1

- ✓ binomial
- ✓ Poisson
- ✓ Bernoulli
  - conditional
  - Expectation

M, V

- ✓ NB(n, p) ✓
- ✓ Uniform Distribution ✓
- ✓ Regression ✓
- ✓ correlation ✓
- ✓ covariance ✓
- ✓ test

Q3. An airline books $\boxed{\text{50 reservation for a plane}}$ with 48 seats. The company assumes 90% of reservation will arrive of the flight.

If this assumption is correct, what is the probability that plane will not accommodate all of the reservation that arrive for the flight?

number of Reservation $= n = 50$

Bernoulli Trials: $\begin{cases} P(\text{reservation}) = 0.9 \\ P(\text{Not reservation}) = 1 - 0.9 = 0.1 \end{cases}$

$n = 50 \quad r \geq 48.$

Using Binomial Distribution $B(n, p)$

$B(50, 0.9)$

$$P[X=r] = {}^nC_r \, p^r q^{n-r}$$

success

$$P[X>48] = P[X=49] + P[X=50]$$

$$= {}^{50}C_{49} (0.9)^{49} (0.1)^{50-49}$$

$$+ {}^{50}C_{50} (0.9)^{50} (0.1)^{50-50}$$

$$= \boxed{0.0304}$$

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

$${}^{50}C_{49} = {}^{50}C_1$$

$$= 50$$

Normal Distribution

Probability Distribution

$\mu = 3$ Per math $\rightarrow$ Poisson distribution $\boxed{0.083}$

Q4. If goals are scored randomly in a game of football at a constant rate of three per match, Calculate the probability that more than 5 goals are scored in a match.

$P[X=x]$ = Poisson Distribution $= \dfrac{e^{-\mu}\mu^{x}}{x!}$

$\mu = 3$ Per math

$P[X > 5 \text{ goals}] = P[X = 6 \text{ goals}] + P[X = 7 \text{ goals}] + P[X = 8 \text{ goals}] + \cdots$

Using PMF

$P[X=0] + P[X=1] + P[X=2] + P[X=3] + P[X=4] + P[X=5] + P[X=6] + \cdots = 1$

Slide 6

$$P(X>5) = 1 - P(X \leq 5)$$

$$P(X>5) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5)]$$

$$P(X>5) = P(X=6) + P(X=7) + P(X=8) + \cdots$$

$$P(X>5) = 1 - \left[\frac{e^{-3}(3)^0}{0!} + \frac{e^{-3}(3)^1}{1!} + \frac{e^{-3}(3)^2}{2!} + \frac{e^{-3}(3)^3}{3!} + \frac{e^{-3}(3)^4}{4!} + \frac{e^{-3}(3)^5}{5!} + \cdots\right]$$

$$\boxed{P(X>5) = 0.0834}$$

$\boxed{\cdot 160}$

Q6. What is the probability of rolling two sixes and three nonsixes in 5 independent casts of a fair die?

What is The Prob 2 Six, 3 Non Sixes.

$\boxed{6}\ \boxed{6}\ \boxed{\bar{6}}\ \boxed{\bar{6}}\ \boxed{\bar{6}}$

SUCCESS — Non Six

$r = 2$

$n = 5$

$B\left(5, \frac{1}{6}\right)$

$P(6) = \frac{1}{6}$ $\quad P(\bar{6}) = \frac{5}{6}$

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

$$= {}^{5}C_{2} \left(\frac{1}{6}\right)^{2} \left(\frac{5}{6}\right)^{5-2} = \underline{0.160}$$

$\boxed{6}\ \boxed{6}\ \boxed{\bar{6}}\ \boxed{\bar{6}}\ \boxed{\bar{6}}$

$6 \quad \bar{6}$

Bernoulli trials (Independent Trials)

Probability Distribution

$\boxed{0.9577}$

(approx.) 0.9

Q7. What is the probability of rolling at most two sixes in 5 independent casts of a fair die?

$n = 5$

$P(X=r) = {}^nC_r\, p^r q^{n-r}$

$p = P(\text{success}) = \frac{1}{6}$

$q = P(\text{failure}) = \frac{5}{6}$

$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$

$= {}^5C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{5-0} + {}^5C_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{5-1} + {}^5C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{5-2}$

$= 0.9577$

$= 0.96 = 0.9$

$\frac{1}{5}$ $\frac{1}{5}$ $\frac{1}{5}$ $\frac{1}{5}$ $\frac{1}{5}$ ✓ $\boxed{0.0204}$

Q8. On a five-question multiple-choice test there are five possible answer, of which one is correct. If a student guesses randomly and independent, what is the probability that she is correct only on questions 1 and 4?

$\boxed{1}$ $\boxed{2}$ $\boxed{3}$ $\boxed{4}$ $\boxed{5}$

$$p \times (1-p) \times (1-p) \times p \times (1-p)$$

$$= p^2(1-p)^3$$

$$P[\text{Question 1 and 4}] = p^2(1-p)^3 = \left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^3 \text{ Ans}$$

S, F — bernoulli Distribution

n - binomial → $P[X=r]$

Probability Distribution

$n=5 \quad r=2 \quad p=\frac{1}{5} \quad q=\frac{4}{5}$

Q9. On a Five-question multiple-choice test there are five possible answer, of which one is correct, if a student guesses randomly and interpedently, what is the probability that she is correct only on $\boxed{\text{two questions?}}$

Using binomial Distribution

$$P[X=2] = {}^5C_2\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^{5-2}$$

$$= \underline{0.2048}$$

Probability Distribution

2 min

$E[X^2] = 154$

Q13. If X is $\boxed{\text{the number of "6"s}}$ that turn up when 72 ordinary dice are independently thrown, find the $E[X^2] = \ldots\ldots$

X is a Discrete Random Variable.

Bernoulli
$P(\text{SUCCESS}) = P(6) = \frac{1}{6}$
$P(\text{failure}) = P(\bar{6}) = 1 - \frac{1}{6} = \frac{5}{6}$

$n = 72$ Dice

$B(n, p) = {}^{n}C_{r}\, p^{r} q^{n-r}$

$V(X) = npq = \cancel{72}^{12} \times \frac{1}{\cancel{6}} \times \frac{5}{\cancel{6}} = 10$

$E[X] = np = \cancel{72} \times \frac{1}{\cancel{6}} = 12$

$E[X^2]$

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

$E[X^2] = \text{Var}(X) + [E[X]]^2$

$= 10 + [12]^2$

$= 10 + 144$

$= 154$

Slide 15

Imp.

$P(B) = 2P(A)$
$P(C) = 2[P(A) + P(B)]$

Q16. A factory makes three different kinds of bolts: Bolt A, Bolt B and Bolt C. The factory produces millions of each bolt every year, but makes twice as many of Bolt B as it does of Bolt A. The number of Bolt C made is twice the total of Bolt A and B combined. Four bolts made by the factory are randomly sampled from all bolts produced by the factory in a given year.

Imp.

Which of the following is most nearly equal to the probability that the sample will contain two of Bolt B and two of Bolt C? 4 bolts

A. $\frac{8}{243}$

B. $\frac{96}{625}$

C. $\frac{384}{2401}$

D. $\frac{32}{243}$

According The question $P(B) = 2P(A) = \frac{2}{9}$

$P(C) = 2[P(A) + P(B)] = \frac{6}{9}$

Using Prob. mass function

$$P(A) + P(B) + P(C) = 1$$

$$P(A) + 2P(A) + 6P(A) = 1$$

$$\Rightarrow 9P(A) = 1$$

$$\left\{ \begin{array}{l} \checkmark\ P(A) = \frac{1}{9} \\ \checkmark\ P(B) = \frac{2}{9} \\ \checkmark\ P(C) = \frac{6}{9} \end{array} \right.$$

4 bolts

bolt A, bolt B, bolt (C)

P(A) P(B) P(C)

0 bolt, 2 bolt, 2 bolt

A B C

Talk

$\underbrace{2 \quad 2}$

bolt B bolt C

0 2 2

Using Trinomial binomial distribution $\Rightarrow \dfrac{n!}{A!\,B!\,C!}\,[B]^{r}\,[C]^{n-r}$

where $n = A+B+C$

$$\frac{n!}{n_1!\,n_2!\,n_3!}\, p^{r} q^{n-r}$$

$$= \frac{4!}{0!\,2!\,2!}\left[\frac{2}{9}\right]^{2}\left[\frac{6}{9}\right]^{4-2} = \frac{4!}{2!\,2!}\left(\frac{2}{9}\right)^{2}\left(\frac{6}{9}\right)^{2} = \frac{32}{243}$$

50% Married
20%
30%] Unmarried

4 People
U/M

Q17. In a large population of people, 50% are married, 20% are divorced, and 30% are single (never married) In a random sample of 4 people, what is the probability that exactly 3 are married?

A. $\binom{4}{3} 0.5^3 \cdot 0.2 \cdot 0.3$

B. $\boxed{\binom{4}{3} 0.5^4}$ ✓

C. 0.5^3

D. 0.5^4

$P(\text{married}) = 0.5$
$P(\text{unmarried}) = 0.5$] Using $B(n, p)$
$= {}^nC_r p^r q^{n-r}$

$P[X=3] = {}^4C_3 (0.5)^3 (0.5)^{4-3}$
$= {}^4C_3 (0.5)^4$

$\binom{4}{3} = {}^4C_3$

$\boxed{\binom{n}{r} = {}^nC_r}$ ✓

Q19. If sum and product of the mean and variance of a binomial distribution are 24 and 128 respectively, then distribution is

$\mu + \sigma^2 = 24$ — (1)

$\mu \cdot \sigma^2 = 128$ — (2)

$B(n, p) = {}^nC_r\, p^r q^{n-r}$

$\mu = np$

$\sigma^2 = npq$

$n = 32$

$\mu = 16 \quad \sigma = 8 \quad p = \frac{1}{2} \quad q = \frac{1}{2}$

A. $\left(\frac{1}{7}+\frac{1}{8}\right)^{12}$

B. $\left(\frac{1}{4}+\frac{3}{4}\right)^{16}$

C. $\left(\frac{1}{6}+\frac{5}{4}\right)^{24}$

D. $\left(\frac{1}{2}+\frac{1}{2}\right)^{32}$ ✓

$(p+q)^n = \left(\frac{1}{2}+\frac{1}{2}\right)^{32}$

1) $= \mu + \sigma^2 = 24$

2) $\mu \sigma^2 = 128$

$$\begin{cases} np + npq = 24 \\ np \cdot npq = 128 \end{cases}$$

$$\left.\begin{aligned} \mu + \sigma^2 &= 24 \\ \mu \sigma^2 &= 128 \end{aligned}\right]$$

16 & 8

$\mu = 16 \quad \sigma^2 = 8$

$np = 16 \quad npq = 8$

$$= \frac{\cancel{np}q}{\cancel{np}} = \frac{8}{16}$$

$$= \underbrace{q = \frac{1}{2} \quad p = \frac{1}{2}}_{} \quad n = 32$$

function $\longrightarrow (p+q)^n \quad 0 \le p \le 1 \quad n = 32$

$$= \left(\frac{1}{2} + \frac{1}{2}\right)^{32}$$

Doyourself

Q20. A family has five children, Assuming that the probability of a girl on a each birth was 0.5 and that the five births were independent, what is the probability the family has at least one girl, given that they have at least one boy?

A. $\frac{31}{32}$

B. $\boxed{\frac{30}{31}}$

C. $\frac{15}{16}$

D. $\frac{5}{31}$

PW

Do yourself

Q21. Let X have a binomial distribution with parameters n and p, Where n is an integer greater than 1 and $0 < p < 1$, if $P(X = 0) = P(X = 1)$, then the value of p is

A. $\frac{1}{n-1}$

B. $\frac{n}{n+1}$

C. $\frac{1}{n+1}$

D. $\frac{1}{1+n^{\frac{1}{n-1}}}$

Probability Distribution

$n=3 \quad p=\frac{1}{6} \qquad B\left(3,\frac{1}{6}\right) = \text{Binomial}$

Q24. In three independent throws of a fair dice, let X denote the number of upper faces showing six. Then the value of $E(3-X)^2$ is.

A. $\frac{20}{3}$

B. $\frac{2}{3}$

C. $\frac{5}{2}$

D. $\frac{5}{12}$

$$E\left[(3-X)^2\right] = E\left[9+X^2-6X\right]$$

$$= 9 + E[X^2] - 6E[X]$$

$$= 9 + \underbrace{Var(X)}_{\substack{\text{variance} \\ npq}} + \{\underbrace{E[X]}_{\substack{\text{mean} \\ np}}\}^2 - 6\underbrace{E[X]}_{\substack{np \\ \text{mean}}}$$

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

$V(X) = npq$

$= \cancel{3} \times \frac{1}{\cancel{6}_2} \times \frac{5}{6}$

$= \frac{5}{12}$

$E[X] = np$

$= \cancel{3} \times \frac{1}{\cancel{6}_2}$

$= \frac{1}{2}$

$$V(X) = \frac{5}{12}$$

$$E\left[(3-X)^2\right] = 9 + \frac{5}{12} + \left(\frac{1}{2}\right)^2 - 6 \times \frac{1}{2}$$

$$= (9-3) + \left[\left(\frac{5}{12}\right) + \frac{1}{4}\right]$$

$$= \frac{20}{3} \checkmark$$

Probability Distribution

$$P(X_i = -1) = P(X_i = +1) = \frac{1}{2}$$ Bernoulli Trials.

Q25. let $X_1, X_2, \ldots X_6$ be independent random variables such that

$\boxed{p(X_i = -1)} = \boxed{P(X_i = 1)} = \frac{1}{2}$, $i = 1, 2, 3, \ldots 6$.

Then $P\left(\boxed{\sum_{i=1}^{6} X_i = 4}\right)$ is

A. $\frac{3}{32}$ ✓

B. $\frac{3}{4}$

C. $\frac{3}{64}$

D. $\frac{3}{16}$

$$P\left[\sum_{i=1}^{6} X_i = 4\right]$$

$$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 = 4$$

$$\boxed{+1 +1 +1 +1} \; [-1] = 4$$

6 Random (5 +) SUCCESS

failure.

$$= {}^6C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^1$$

$$= \frac{{}^6C_5}{2^6} = \frac{6}{64} = \frac{3}{32}$$

Slide 27

Do Yourself

Q27. Six identical fair dice are thrown independently . Let S denote the number of dice showing even number on their upper faces. Then the variance of the random variable S is

A. $\frac{1}{2}$

B. 1

C. $\frac{3}{2}$

D. 3

Do yourself

(B) $\frac{119}{128}$

Q28. If X–B in $\left(8, \frac{1}{2}\right)$, then $P\ (|X-4| \leq 2)$ is

A. $\frac{9}{128}$

B. $\frac{119}{128}$

C. $\frac{91}{128}$

D. $\frac{109}{128}$

Do yourself

Q30. Let the random variable X – Bin (5, p) such and P (X = 2) =2p (X – 3).

Then the variance of x is

A. $\frac{10}{3}$

B. $\frac{10}{9}$

C. $\frac{5}{3}$

D. $\frac{5}{9}$

THANK - YOU