SWAYAM NPTEL COURSE ON

MINE AUTOMATION AND DATA ANALYTICS

By

Prof. Radhakanta Koner
Department of Mining Engineering
Indian Institute of Technology (ISM) Dhanbad

Module 8: Inferential Statistics

Lecture 19B: Discrete Random Variable Part II

CONCEPTS COVERED

In Continuation:

- Properties of Variance and Standard Deviation.
- Bernoulli Distribution
- Binomial Distribution
- Uniform Distribution

Properties of Variance

Let X be a random variable, let c be a constant, then

$$Var(cx) = c^2 Var(X)$$
$$Var(X + c) = Var(X)$$

If a and b are constants, $V(aX + b) = a^2V(X)$

<u>Proof.</u>

We know $E(ax + b) = a\mu + b,$
$Var(ax + b) = E(ax + b - a\mu - b)^2 = a^2E(X - \mu)^2 = a^2 Var(X)$

JAN 2024

Variance of sum of two random variables

- The expected value of the sum of random variables equals the sum of the individual expected values. In other words, let X and Y be two random variables. Then,

$$E(X+Y) = E(X) + E(Y)$$

- Where as $Var(X + X) = Var(2X) = 4Var(X) \neq Var(X) + Var(X)$
- $4Var(X) \neq 2\,Var(X)$

- Is this statement always true in all cases?

Independent random variables

Definition

Random variables X and Y are considered independent if the knowledge of the value of one of them does not alter the probabilities associated with the other.

Example: Roll a dice twice. S = $\{(1,1), \ldots, (6,6)\}$

- **X = the outcome of the first dice.**
- **Y = the outcome of the second dice**
- **Knowing X = i does not change the probability of Y taking any value of 1,2,..,6 .**
- **X and Y are independent random variables.**

Variance of sum of independent random variables

Result

Let X and Y be independent random variables. Then

$$Var(X + Y) = Var(X) + Var(Y)$$

Example: Rolling a dice twice

- Let *X* represent the outcome of one fair dice, and let *Y* represent the outcome of another fair dice.
- We observe that $E(X) = E(Y) = 3.5$.
- The sum of outcomes of both dice rolled together, denoted as $X+Y$, is calculated to have

  Var(X) = Var(Y) = 2.917.
- Since *X* and *Y* are independent, we compute Var(X+Y)=2.917+2.917=5.83 , which aligns with the result obtained by applying the computational formula.

Variance of sum of many independent random variables

- The outcome stating that the variance of the sum of independent random variables equals the sum of the variances applies not only to two but to any number of random variables.
- Let $X_1, X_2, \ldots, X_k$ be k discrete random variables. Then,

$$\boxed{Var\left(\sum_{i=1}^{k} Xi\right) = \sum_{i=1}^{k} Var(Xi)}$$

JAN 2024

Standard Deviation (SD) of a random variable

Definition: The quantity SD(X) = $\sqrt{Var(X)}$ is the standard deviation of X.

Therefore, the standard deviation (SD) is defined as the positive square root of the variance.

Similar to the expected value, the standard deviation is expressed in the same units as the random variable.

Properties of Standard Deviation

- **Let X be a random variable, let c be a constant, then**

$$\boldsymbol{SD(cX) = c.SD(X)}$$
$$\boldsymbol{SD(X + C) \; = \; SD(X)}$$

Binomial Distribution

Bernoulli Trail

- A trial or experiment, where the outcome can be categorized as either a success or a failure, is referred to as a Bernoulli trial.
- The sample space S = {Success, Failure}
- Let X represent a random variable that takes the value 1 if the outcome is a success and 0 if the outcome is a failure.
- X is referred to as a Bernoulli random variable.

Examples of Bernoulli trails

Experiment 1:

Tossing a coin: S = {Head, Tail}

Success : Head

Failure : Tail

Experiment 2:

Rolling a dice: S = {1,2,3,4,5,6}

Success : Getting a six.

Failure : Getting any other number

Non Bernoulli trail

- Experiment: Selecting a person at random and inquiring about their age.

- This experiment does not qualify as a Bernoulli trial because it does not entail only two possible outcomes.

Bernoulli random variable

- **A random variable that can assume either the value 1 or 0 is termed a Bernoulli random variable.**
- **Let X be a Bernoulli random variable that takes on the value 1 with probability p.**
- **The probability distribution of this random variable is as follows:**

X	0	1
$P(X = xi)$	$1 - p$	p

- **Expected value of a Bernoulli random variable:**

$$E(X) = \sum_{i=1}^{\infty} x_i P(x_i)$$

E(X) = 0 x (1- p) + 1 x p = p

- **Variance of a Bernoulli random variable : $V(X) = p - p^2 = p(1-p)$**

Variance of Bernoulli Distribution

- The maximum variance occurs when $p = \frac{1}{2}$, when success and failure are equally probable.
- In simpler terms, the most uncertain Bernoulli trials, characterized by the largest variance, resemble the tossing of a fair coin.

Independent and Identically distributed Bernoulli trails

- N = 3 independent trials
- Let n = 3 independent bernoulli trails.
- Let p is the probability of success.
- The probability of outcomes of the independent trials are

$Sl.No$	$Outcome$	$Number\ of\ successes$	$Probabilities$
1	(s,s,s)	3	$p.p.p$
2	(s,s,f)	2	$p.p.(1-p)$
3	(s,f,s)	2	$p.(1-p).p$
4	(s,f,f)	1	$p.(1-p).(1-p)$
5	(f,s,s)	2	$(1-p).p.p.$
6	(f,s,f)	1	$(1-p).p.(1-p)$
7	(f,f,s)	1	$(1-p).(1-p).p$
8	(f,f,f)	0	$(1-p).(1-p).(1-p)$

N = 3 independent trails, X = number of successes

- Let n = 3 independent Bernoulli trails
- Let p is the probability of success.
- Let X = number of successes in 3 independent trails.
- The probability distribution of X

X	0	1	2	3
$P(X = i)$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3

N independent trails , X = number of successes

- Consider any outcome that results in a total of i successes.
- The outcome will have a total of i successes and (n- i) failures.
- Probability of i success and (n- i) failures = $p^i.(1-p)^{n-i}$
- There number of different outcomes that result in i successes and (n- i) failures = $\binom{n}{i}$
- The probability of i success in n trials is given by

$$P(X=i) \ = \binom{n}{i}.pi.(1-p)^{n-i}$$

Binomial Random Variable

Definition

- **Let X be a binomial random variable with parameters n and p, which denotes the number of successes in n independent Bernoulli trials, where each trial has a success probability of p**
- **X takes values 0,1, 2,3,…,n with the probability.**

$$P(X = i) = \binom{n}{i}.pi.(1-p)^{n-i}$$

Example: Tossing a coin thrice

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

- **Success = head, Failure = tails**
- **X is the random variable which counts the number of heads in the tosses.**
- **N = 3, P = 0.5**
- **Probability Mass function**

X	0	1	2	3
$P(X=x_i)$	$\binom{3}{0}(1/2)^0(1/2)^3$ = 1/8	$\binom{3}{1}(1/2)^1(1/2)^2$ = 3/8	$\binom{3}{2}(1/2)^2(1/2)^1$ = 3/8	$\binom{3}{3}(1/2)^3(1/2)^0$ = 1/8

Shape of the pmf for same n different p

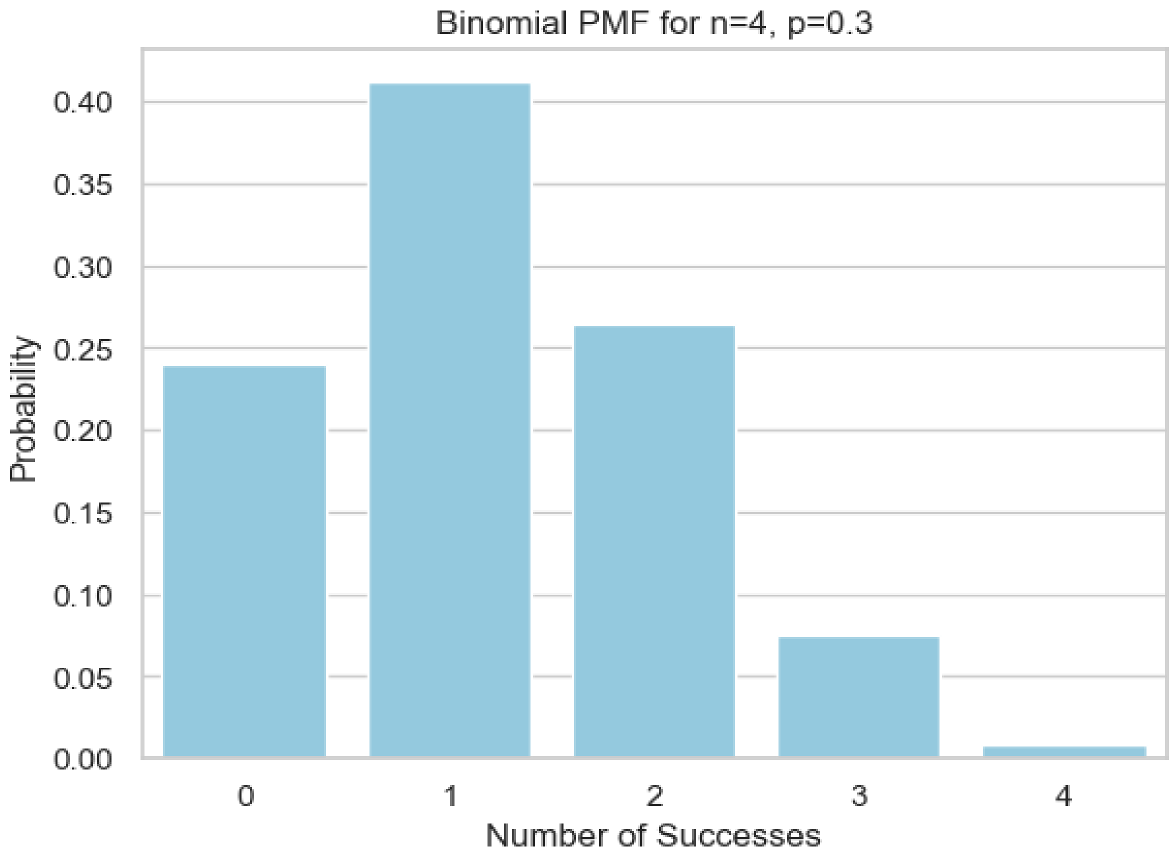
A binomial distribution is

- Right skewed if $p < 0.5$
- symmetric if $p = 0.5$
- Left skewed if $p > 0.5$

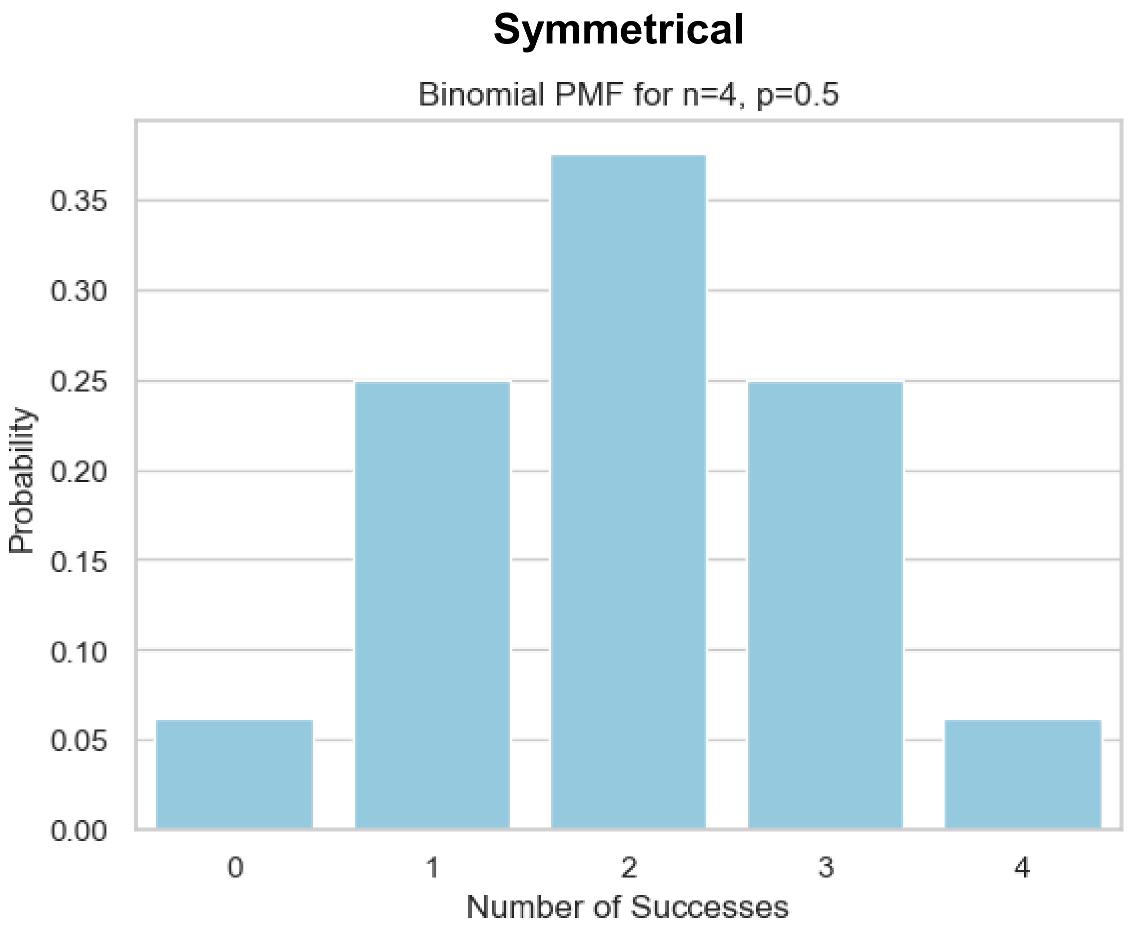
- We demonstrate the same for n = 4 and different p

N = 4 , p = 0.3 , X = number of success

Right Skewed

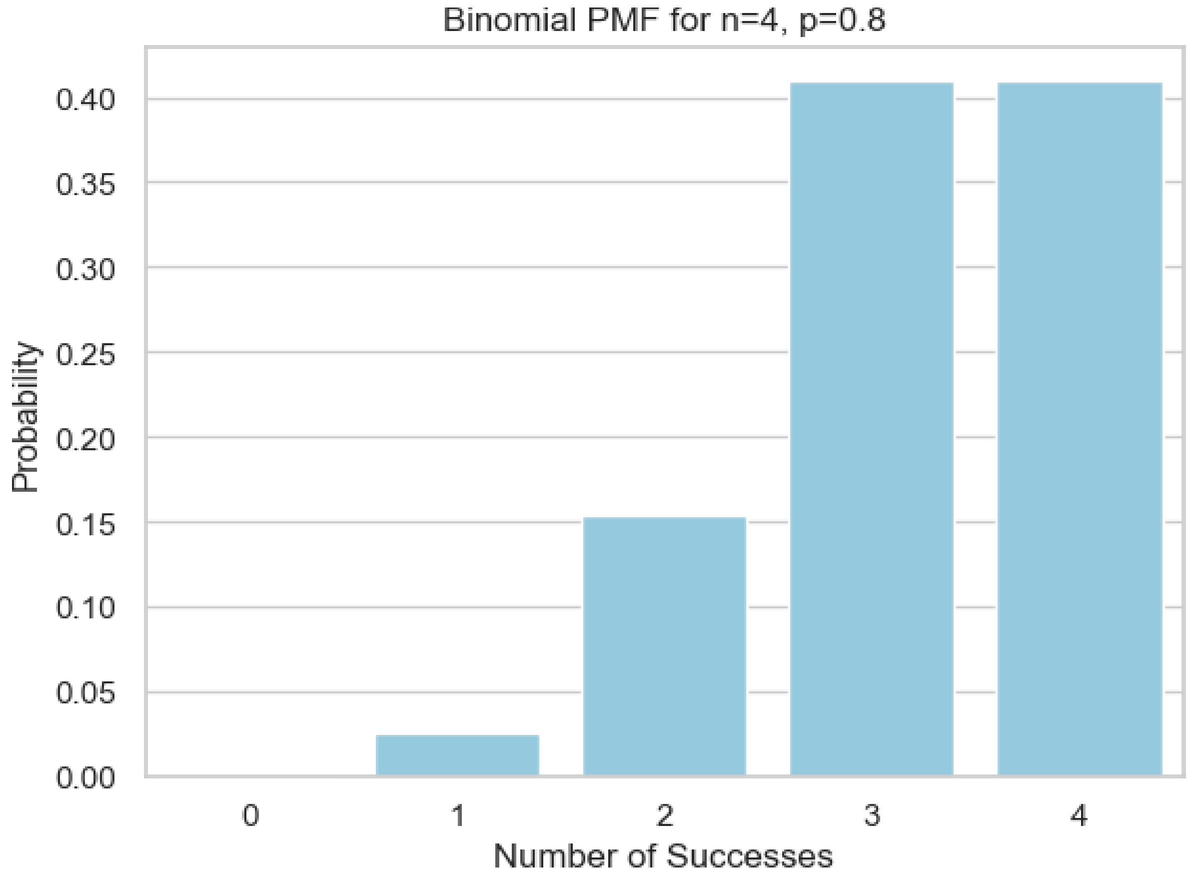


N = 4 , p = 0.5 , X = number of success

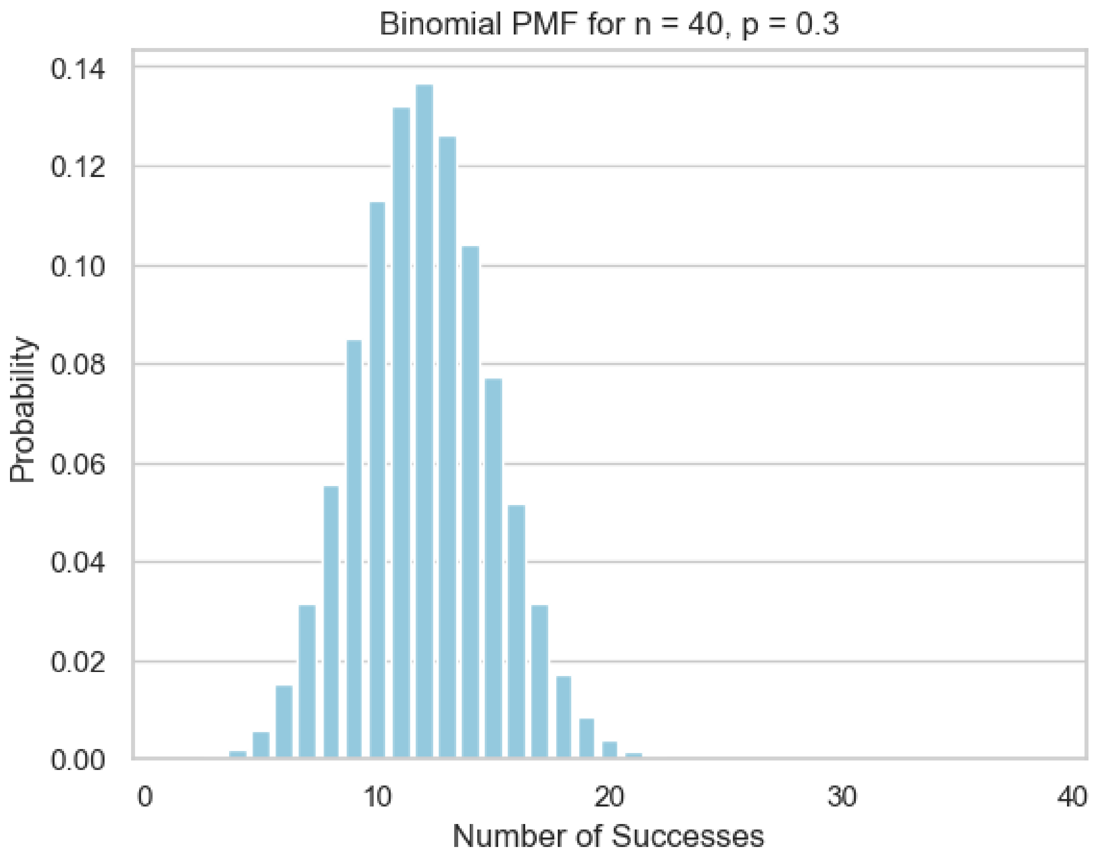


N = 4 , p = 0.8 , X = number of success

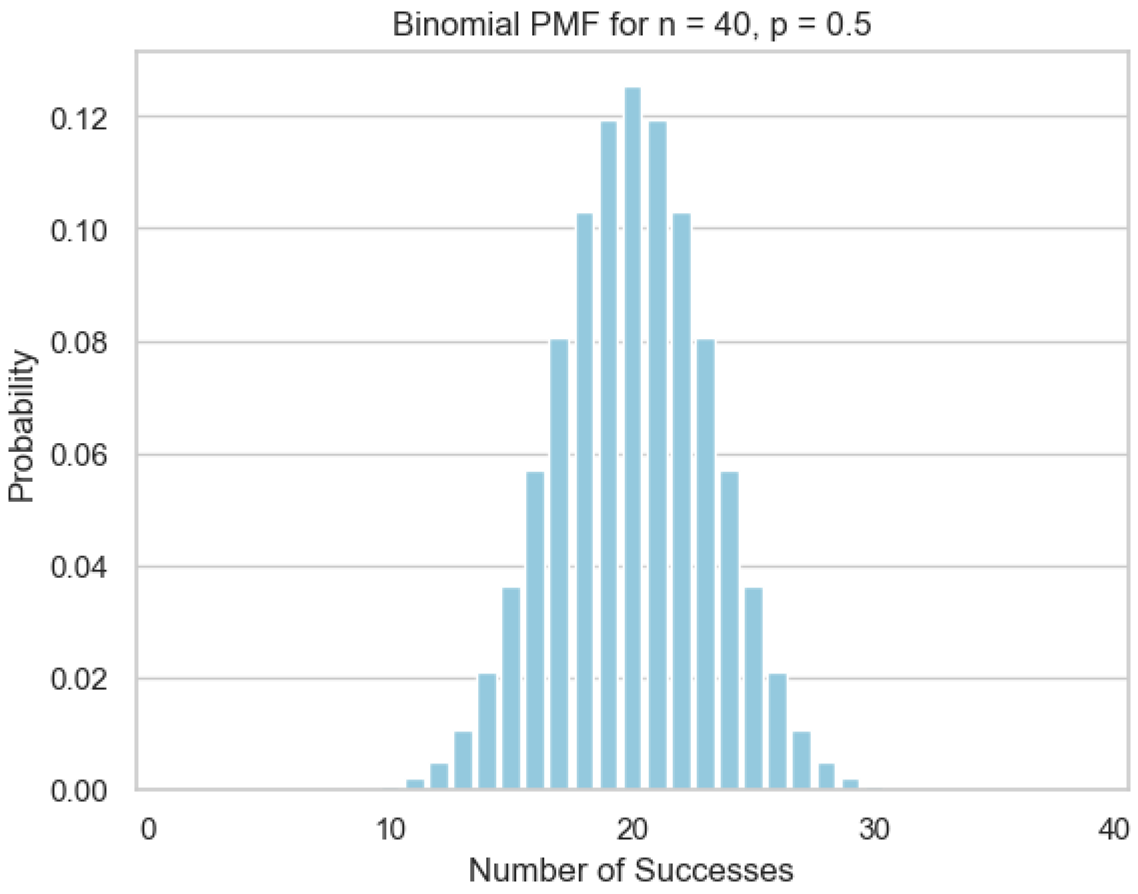
Left Skewed



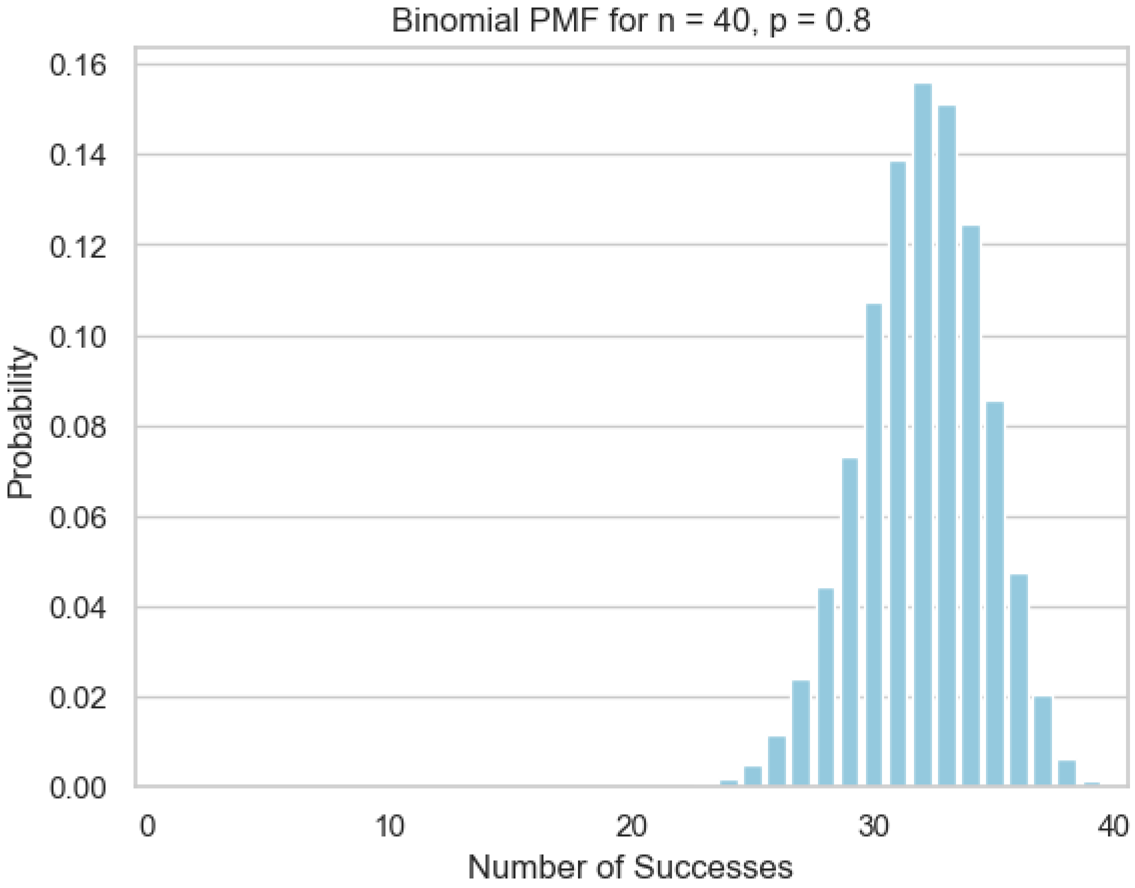
Graph of pmf of Binomial distribution
Right Skewed – for larger n



Graph of pmf of Binomial distribution Symmetrical – for larger n



Graph of pmf of Binomial distribution left skewed – for larger n



Effect of n and p on the shape of the distribution

- For small n and small p, the distribution is right-skewed.
- For small n and large p, the distribution is left-skewed.
- For small n and $p=0.5$, the distribution is symmetric.
- As n becomes large, the binomial distribution tends towards symmetry.

Expectation and Variance of Binomial Random Variable

- A binomial random variable X ~ Bin (n, p) equals the number of successes in n independent trials when each trial is a success with probability p.
- We can represent X as

$$X = X_1 + X_2 + \cdots . + X_n$$

- Where X_i is equal to 1 if trail i is a success and is equal to 0 if trail i is a failure.

$$P(X_i = 1) = p$$

$$P(Xi = 0) = 1 - p$$

Expectation and Variance of Binomial Random Variable

$X = X1 + X2 + \cdots + Xn$

- $E(X) = E(X1 + X2 + \cdots + Xn) = E(X1) + E(X2) + \ldots + E(Xn)$
  $= p + p + \ldots .. + p = np$

- $V(X) = V(X1 + X2 + \cdots + Xn) = V(X1) + V(X2) + \ldots + V(Xn)$
  $= p(1-p) + p(1-p) + \ldots .. + p(1-p) = p(1-p)$

Result:

- **The expectation of a binomial random variable :** $E(X) = np$
- **The variance of a binomial random variable :** $V(X) = np(1-p)$

Example: Tossing a coin 500 times

- If a fair coin is tossed 5000 times, what is the standard deviation of the number of times a head appears?
- Let X = the number of heads in 5000 fair coin tosses. Then $X \sim Bin(500, 1/2)$
- $E(X) = np = (5000)(1/2) = 2500$
- $V(X) = np(1-p) = 5000(1/2)(1-(1/2)) = 1250$
- $S(X) = \sqrt{V(X)} = 35.35$

Finding probability given expectation and n

In a series of 10 coin tosses, the expected number of heads is 6. What is the probability of obtaining 8 heads?

Given E(X) = 6

- **We already know the probability of getting a fair head each in 10 independent coin tosses = ½**
- **But we don't know whether it is a fair coin.**
- $np = 6,\ p = 0.6$
- $Bin(n = 10, p = 1/6)$
- $P(X = 8)$?

$$P(X = i) = \binom{n}{i} . pi.(1-p)^{n-i}$$

$$= \binom{10}{8} . (0.6)^8 . (1 - 0.6)^2 = 0.121$$

Uniform Random Variable

- A discrete uniform random variable is a type of random variable that characterizes outcomes from a finite set of equally probable values.
- In essence, it indicates a scenario where each potential outcome has an identical likelihood of occurrence.
- The term "uniform" is employed because the probabilities are uniformly distributed across the range of possible values.

Uniform Random Variable

- Let X be a random variable that is equally likely to take any of the values 1,2,3,...,n
- The probability mass function (PMF) of a discrete uniform random variable is given by

X	1	2	...	n
P(X = x_i)	1/n	1/n	...	1/n

- $E(X) = (1 \times 1/n) + (2 \times 1/n) + \ldots. + (n \times 1/n) = (n+1)/2$
- $E(X^2) = (1 \times 1/n) + (4 \times 1/n) + \ldots. + (n^2 \times 1/n) = (n+1)(2n+1)/6$
- $Var(X) = E(X^2) - (E(X))^2 = (n^2-1)/12$

JAN 2024

Example of a Uniform random variable

Example 1

- Let's take a fair six-sided die as an example. The potential outcomes when rolling the die are 1, 2, 3, 4, 5, and 6.
- Because each face has an equal probability of landing face up, the random variable representing the outcome of the die roll adheres to a discrete uniform distribution with n=6.
- The probability of obtaining any specific number (such as 3) is 1/6

Example 2

- Another instance is selecting a card randomly from a thoroughly shuffled standard deck of 52 cards, where each card carries an equal probability of 1/52

REFERENCES

- Introduction to Probability and Statistics for Engineers and Scientists, Sixth Edition, Sheldon M. Ross
- Statistical Methods Combined Edition (Volume I& II), N G Das

CONCLUSION

- We have discussed the properties of Variance and Standard Deviation.
- We have discussed about Bernoulli Distribution
- We have discussed Binomial Distribution
- We have discussed about Uniform Distribution


THANK YOU