Bernoulli Trial

- A trial is said to be Bernoulli trial if it can results in a success or a failure.

Here is a simple example of a Bernoulli trial. From a standard deck of cards, you pick a card, note whether it is a club or not. So the outcome of the trial can be classified in two categories: selecting a club (success) and selecting another suit (failure).

The probabilities of success and failure are denoted by "p" and "q" respectively. If the random variable X represents the number of clubs selected, then the possible values of the random variable are 0 and 1.

Outcome		S or F?
		F
	Club	S

> The random variable "X" representing the number of successes in Bernoulli trials is called a Bernoulli random variable

Bernoulli distribution

The probability distribution of the Bernoulli variable "X" is called as Bernoulli distribution.

The probability mass function of Bernoulli distribution is given below:

$$P(X=x)=f(x)=\begin{cases} p^x q^{1-x} & ;x=0,1 \\ 0 & ;otherwise \end{cases}$$

Where

- x = number of success
- p = probability of success
- q = probability of failure

Note: "p" is the **parameters** of the Bernoulli distribution

Historical Note

Bernoulli random variable is named in honor of the mathematician Jacob Bernoulli (1654-1705).

Mean, Variance and S.D of Bernoulli distribution

Measure	Formula
Mean	$\mu = p$
Variance	$\sigma^2 = pq$
Standard Deviation	$\sigma = \sqrt{pq}$

The prefix "bi" means "two". This should help you to remind that binomial experiments deal with situations in which there are only two outcomes i.e. success and failure.

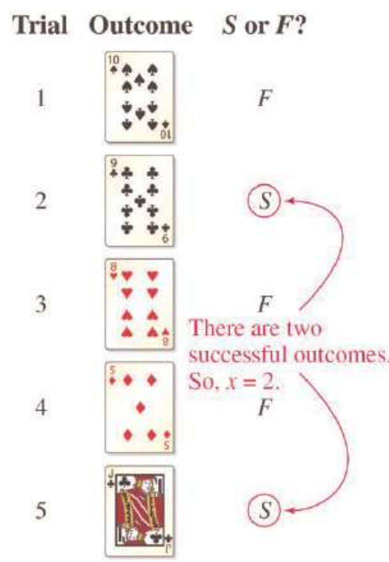
Binomial experiment

An experiment that has the following properties is called Binomial experiment:

- Every trial results in a success or a failure.
- The successive trials are independent.
- The probability of successes remains constant from trial to trail.
- The number of trials is fixed in advance.

Here is a simple example of a binomial experiment. From a standard deck of cards, you draw 5 cards in succession, note whether it is a club or not, and **replace** the card. So the outcomes of each trial can be classified in two categories: selecting a club (success) and selecting another suit (failure).

The probabilities of success and failure are denoted by "p" and "q" respectively. The **probability of success remains the same** because the card once drawn has been **replaced** before the next draw. If the random variable X represents the number of clubs selected, then the possible values of the random variable are 0, 1, 2, 3, 4, and 5. Note that X is a discrete random variable because its possible values can be listed.



The random variable "X" representing the number of successes in a binomial experiment is called a binomial random variable

Binomial Distribution

The probability distribution of the binomial variable "X" is called as binomial distribution.

The probability mass function of binomial distribution is given below:

$$P(X=x)=f(x)=\begin{cases}\binom{n}{x}p^x q^{n-x} & ;x=0,1,2,\dots,n\\ \\ 0 & ;otherwise\end{cases}$$

Where

- x = number of success
- p = probability of success
- q = probability of failure
- n = number of trials that are fixed in advance

Note: "n" and "p" are the **parameters** of the binomial distribution.

Historical Note

The binomial distribution is a very important discrete probability distribution. It was discovered by James Bernoulli about the year 1700.

The terms "success" and "failure" are used in the binomial doesn't necessarily mean a success is good and failure is bad. Success mean that you get the outcome you want to count, and failure means you get the outcome you don't want to count. For example, if you select ten 18-year-old male drivers, then a success may be an 18-year-old driver who was involved in an accident.

Mean, Variance and S.D of Binomial Distribution

Measure	Formula
Mean	$\mu = np$
Variance	$\sigma^2 = npq$
Standard Deviation	$\sigma = \sqrt{npq}$

EXAMPLE 8.01

Find complete binomial distribution having n = 5 and p = 1/2

Solution Here $n=5 \quad \Rightarrow x=0,1,2,3,4,5$

And $p=1/2 \quad \Rightarrow q=1-p=1-1/2=1/2$

Now $f(x)=\binom{n}{x}p^x q^{n-x}$

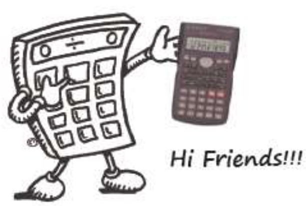
$$\Rightarrow f(x)=\binom{5}{x}(1/2)^x(1/2)^{5-x}$$

Hence the complete binomial distribution becomes:

REMEMBER

A Binomial distribution having n = 5 and p = 1/2 can also be written as $\left(\frac{1}{2}+\frac{1}{2}\right)^5$

x	$f(x)=\binom{5}{x}(1/2)^x(1/2)^{5-x}$
0	$f(0)=\binom{5}{0}(1/2)^0(1/2)^{5-0}=(1)(1/2)^5=1/32$
1	$f(1)=\binom{5}{1}(1/2)^1(1/2)^{5-1}=(5)(1/2)^5=5/32$
2	$f(2)=\binom{5}{2}(1/2)^2(1/2)^{5-2}=(10)(1/2)^5=10/32$
3	$f(3)=\binom{5}{3}(1/2)^3(1/2)^{5-3}=(10)(1/2)^5=10/32$
4	$f(4)=\binom{5}{4}(1/2)^4(1/2)^{5-4}=(5)(1/2)^5=5/32$
5	$f(x)=\binom{5}{5}(1/2)^5(1/2)^{5-5}=(1)(1/2)^5=1/32$



EXAMPLE 8.02

An event has the p = 3/8 and n = 5 find the probability of:

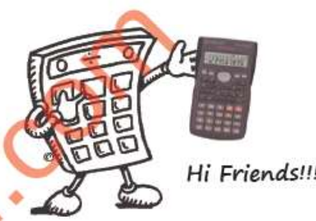
(i) $P(X=3)$ (ii) $P(X>3)$ (iii) $P(X\geq 3)$ (iv) $P(X\leq 3)$

Solution Here $n=5 \quad \Rightarrow x=0,1,2,3,4,5$

And $p=3/8 \quad \Rightarrow q=1-p=1-3/8=5/8$

Now $P(X=x)=\binom{n}{x}p^x q^{n-x}$

$$\Rightarrow P(X=x)=\binom{5}{x}(3/8)^x(5/8)^{5-x}$$



(i) $P(X=3)$

$$P(X=3)=\binom{5}{3}(3/8)^3(5/8)^{5-3}$$

$$=(10)(3/8)^3(5/8)^2=0.21$$

In binomial distribution we can not find the probability of the form $P(x=2.3)$ because the binomial r.v. X can take only the integer values.

(ii) $P(X>3)$

$$P(X>3)=P(X=4)+P(X=5)$$

$$=\binom{5}{4}(3/8)^4(5/8)^{5-4}+\binom{5}{5}(3/8)^5(5/8)^{5-5}$$

$$=(5)(3/8)^4(5/8)^{5-4}+(1)(3/8)^5(5/8)^{5-5}$$

$$=(5)(3/8)^4(5/8)+(1)(3/8)^5=0.07$$

(iii) $P(X\geq 3)$

$$P(X\geq 3)=P(X=3)+P(X=4)+P(X=5)$$

$$=\binom{5}{3}(3/8)^3(5/8)^{5-3}+\binom{5}{4}(3/8)^4(5/8)^{5-4}+\binom{5}{5}(3/8)^5(5/8)^{5-5}$$

$$=(10)(3/8)^3(5/8)^{5-3}+(5)(3/8)^4(5/8)^{5-4}+(1)(3/8)^5(5/8)^{5-5}$$

$$=(10)(3/8)^3(5/8)^2+(5)(3/8)^4(5/8)+(1)(3/8)^5=0.28$$

(iv) $P(X \le 3)$

$$P(X \le 3) = 1 - P(X > 3)$$

$$= 1 - 0.07 = 0.93$$

TIP

$$P(X \le a) + P(X > a) = 1$$

EXAMPLE 8.03

Find mean, variance and S.D for a binomial distribution having n = 20 and p = 0.3

Solution We know that for a binomial distribution

$$Mean = np$$
$$= (20)(0.3) = 6 \qquad (q = 1 - p = 1 - 0.3 = 0.7)$$

$$Variance = npq$$
$$= (20)(0.3)(0.7) = 4.2$$

$$S.D = \sqrt{npq}$$
$$= \sqrt{(20)(0.3)(0.7)} = 2.05$$

REMEMBER

For a binomial distribution
Mean > Variance

EXAMPLE 8.04

The mean and variance of a binomial distribution are 42 and 12.6. Find the values of the parameters "n" and "p"?

Solution Given that

$$Mean = 40$$

$$\Rightarrow np = 40 - - - - - (a)$$

And $Variance = 12.6$

$$\Rightarrow npq = 12.6 - - - - - (b)$$

Putting $np = 40$ in equation (b) we have:

$$4.2q = 12.6 \Rightarrow q = \frac{12.6}{4.2} = 0.3$$

Now since $p = 1 - q \Rightarrow p = 1 - 0.3 = 0.7$

Putting $p = 0.7$ in equation (b) we have:

$$n(0.7) = 42 \Rightarrow n = \frac{42}{0.7} = 60$$

Hence the values of the parameters are:

n = 60 and p = 0.7

Prove that the mean of the Binomial distribution is "np"

Proof:

We know that $Mean = \mu = E(X) = \sum_{x=0}^{n} x f(x)$

$$\Rightarrow Mean = \mu = E(X) = \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x q^{n-x} \qquad \left(\because f(x) = \binom{n}{x} p^x q^{n-x} \right)$$

$$= 0\binom{n}{0} p^0 q^{n-0} + 1\binom{n}{1} p^1 q^{n-1} + 2\binom{n}{2} p^2 q^{n-2} + \ldots + n\binom{n}{n} p^n q^{n-n}$$

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

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

$$= np\ \left[q + p \right]^{n-1}$$

μ (meu)

$= np$ Hence proved

Properties of Binomial Distribution

- The mean and variance of the binomial distribution are: "*np*" and "*npq*" respectively
- For the binomial distribution *Mean > Variance*
- The shape of the binomial distribution depends on the values of "*n*" and "*p*"
- The Binomial distribution is
  - Symmetric if $p = q = 1/2$
  - Positively skewed if $p < q$
  - Negatively skewed if $p > q$
- The Binomial distribution approach to Normal distribution; as $n \to \infty$ such that $np > 5$ and $nq > 5$
- The moments about mean of the binomial distribution are:
  - $\mu_1 = 0$
  - $\mu_2 = npq$
  - $\mu_3 = npq(1-2p)$
  - $\mu_4 = 3n^2p^2q^2 + npq(1-6pq)$

- For the binomial distribution:
  - $coefficient\ of\ variation = \sqrt{\frac{q}{np}} \times 100$
  - $coefficient\ of\ skewness = \frac{q-p}{\sqrt{npq}}$
  - $coefficient\ of\ kurtosis = 3 + \frac{1-6pq}{npq}$

EXAMPLE 8.05

Is it possible to have a binomial distribution with mean = 5 and S.D = 3?

Solution Given that

$$Mean = 5$$
$$S.D = 3 \Rightarrow Variance = 3^2 = 9$$

But we know that for any binomial distribution Mean > Variance

Hence it is not possible to have a binomial distribution with mean = 5 and S.D = 3.

EXAMPLE 8.06

If X is a binomial r.v with n = 20 and p = 0.5 then find:

(i) Coefficient of variation
(ii) Coefficient of skewness
(iii) Coefficient of kurtosis

Solution **(i) Coefficient of variation**

$$Coefficient\ of\ variation = \sqrt{\frac{q}{np}} \times 100$$
$$\Rightarrow Coefficient\ of\ variation = \sqrt{\frac{0.5}{(20)(0.5)}} \times 100 = 22.4\% \qquad (q = 1 - p = 1 - 0.5 = 0.5)$$

(ii) Coefficient of skewness

$$Coefficient\ of\ skewness = \frac{q-p}{\sqrt{npq}}$$

$$\Rightarrow Coefficient\ of\ skewness = \frac{0.5-0.5}{\sqrt{(20)(0.5)(0.5)}} = 0$$

Hence the distribution is **symmetric**.

(iii) Coefficient of kurtosis

$$Coefficient\ of\ kurtosis = 3 + \frac{1-6pq}{npq}$$

$$\Rightarrow Coefficient\ of\ kurtosis = 3 + \frac{1-6(0.5)(0.5)}{(20)(0.5)(0.5)} = 2.9$$

Hence the distribution is **platykurtic**.

EXAMPLE 8.07

If X is a binomial r.v with n = 20 and p = 0.3 then find E(2X-3)?

Solution We know that for a binomial distribution

$$\begin{aligned} Mean = E(X) &= np \\ &= (20)(0.3) \\ &= 6 \end{aligned}$$

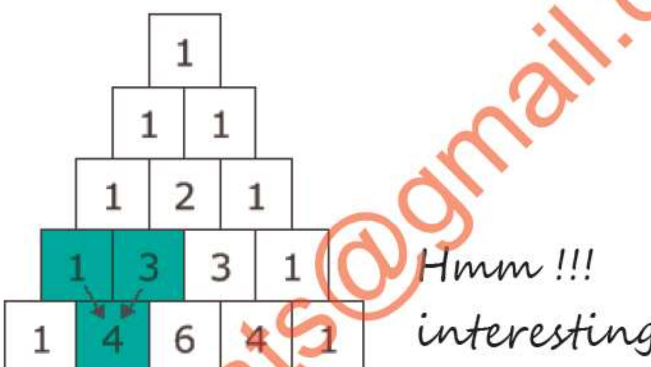
$$\begin{aligned} E(2X-3) &= 2E(X) - 3 \\ &= 2(6) - 3 \\ &= 9 \end{aligned}$$

Pascal's Triangle!!!

Consider the binomial expansion:

$$(p+q)^n = \binom{n}{0} p^n q^0 + \binom{n}{1} p^{n-1} q^1 + \binom{n}{2} p^{n-2} q^2 + \cdots\cdots + \binom{n}{n} p^o q^n$$

The coefficients $\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots\ldots, \binom{n}{n}$ are called the binomial coefficients. These coefficients can be easily written down by using an arrangement of numbers, called Pascal's triangle given below:



and so on . . .

Historical Note

In 1653, Blaise Pascal created a triangle of numbers called Pascal's triangle that can be used in the binomial distribution

Power	Binomial Expansions	Coefficients
2	$(p+q)^2 = p^2 + 2pq + q^2$	1 2 1
3	$(p+q)^3 = p^3 + 3p^2q + 3pq^2 + q^3$	1 3 3 1
4	$(p+q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$	1 4 6 4 1
and so on…		

EXAMPLE 8.08

Expand the binomial distribution $\left(\frac{1}{3}+\frac{2}{3}\right)^4$

Solution Here p = 1/3 and q = 2/3 and n = 4

Now since $(p+q)^4 = p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4$

$$\therefore \left(\frac{1}{3}+\frac{2}{3}\right)^4 = \left(\frac{1}{3}\right)^4 + 4\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right) + 6\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 + 4\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^3 + \left(\frac{2}{3}\right)^4$$

$$\Rightarrow \left(\frac{1}{3}+\frac{2}{3}\right)^4 = \frac{1}{81}+\frac{8}{81}+\frac{24}{81}+\frac{32}{81}+\frac{16}{81}$$

This is the required expansion.

Test Yourself

1) Find complete binomial distribution having n = 6 and p = 1/4
2) Find mean, variance and S.D for a binomial distribution having n = 15 and p = 0.7
3) The mean and variance of a binomial distribution are 1.2 and 0.84. Find the values of the parameters "n" and "p"?
4) If X is a binomial r.v with n = 14 and p = 0.8 then find:

    (i) Coefficient of variation
    (ii) Coefficient of skewness
    (iii) Coefficient of kurtosis

5) An event has the p = 1/8 and n = 5 find the probability of:

    (i) $P(X=2)$ (ii) $P(X>2)$ (iii) $P(X\geq 4)$ (iv) $P(X\leq 2)$

6) Expand the binomial distribution $\left(\frac{1}{4}+\frac{3}{4}\right)^5$