Unit 3

The Bernoulli Process

the Bernoulli process must possess the following properties:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Bernoulli Distribution.

A random variable X which takes two values 0 and 1 with probabilities q and p respectively

i.e . P (X= 1) = p

P (X= 0) = 1-p =q

is called a Bernoulli variate and is said to have a Bernoulli distribution.

Examples of Bernoulli trials

- Experiment: Tossing a coin: $S = \{Head, Tail\}$
  - Success: Head
  - Failure: Tail
- Experiment: Rolling a dice: $S = \{1, 2, 3, 4, 5, 6\}$
  - Success: Getting a six.
  - Failure: Getting any other number.
- Experiment: Opinion polls: $S = \{Yes, No\}$
  - Success: Yes
  - Failure: No
- Experiment: Salesperson selling an object: $S = \{Sale, No\ sale\}$
  - Success: Sale
  - Failure: No sale
- Experiment: Testing effectiveness of a drug: $S = \{Effective, Not\ effective\}$

Binomial Experiment

- A <u>binomial experiment</u> has the following properties:
  - ❑experiment consists of **n identical and independent trials**
  - ❑each trial results in one of two outcomes: success or failure
  - ❑Probabilities for all trials
    - P(success) = p
    - P(failure) = q (= 1 – p)
  - ❑The random variable of interest, X, is the number of successes in the n trials.

Binomial Distribution A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

or

Definition. *A random variable X is said to follow binomial distribution if it assumes only non-negative values and its probability mass function is given by*

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

OR

Probability density function

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

- X has a binomial distribution with parameters n and p
- The mean (expected value): np
- Variance: npq

Relation between Mean and Variance

$$\frac{Mean}{Variance}=\frac{np}{npq}=\frac{1}{q}>1$$

$$Mean > Variance$$

Remarks

1. Binomial distribution is a legitimate probability distribution since

$$(q+p)^n=\binom{n}{0}q^n+\binom{n}{1}pq^{n-1}+\binom{n}{2}p^2q^{n-2}+\cdots+\binom{n}{n}p^n$$
$$=b(0;n,p)+b(1;n,p)+b(2;n,p)+\cdots+b(n;n,p).$$

Since $p+q=1$, we see that

$$\sum_{x=0}^{n} b(x;n,p)=1,$$

Or

$$\sum_{r=0}^{n} P(X=r)=\sum_{r=0}^{n} nC_r\, q^{n-r}\, p^r$$
$$=(q+p)^n=1$$

(2) *The name 'binomial distribution' is given since the probabilities* $nC_r\, q^{n-r}\, p^r$ $(r = 0, 1, 2, \ldots, n)$ *are the successive terms in the expansion of the binomial expression* $(q + p)^n$.

(3) *If we assume that n trials constitute a set and if we consider N sets, the frequency function of the binomial distribution is given by* $f(r) = N\, p(r) = N \cdot nC_r\, q^{n-r}\, p^r$, $r = 0, 1, 2, \ldots, n$. *In other words, the number of sets in which we get exactly r successes (the occurrences of the event A)* $= N \cdot nC_r\, q^{n-r}\, p^r$; $r = 0, 1, 2, \ldots, n$.

Example 1

The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Solution:-

Assuming that the tests are independent and p = 3/ 4 for each of the 4 tests, we obtain

$$b\left(2; 4, \frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \left(\frac{4!}{2!\,2!}\right)\left(\frac{3^2}{4^4}\right) = \frac{27}{128}.$$

Example 2

Example 7·1. *Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.*

Solution. p = Probability of getting a head = $\frac{1}{2}$

q = Probability of not getting a head = $\frac{1}{2}$

The probability of getting x heads in a random throw of 10 coins is

$$p(x) = \binom{10}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{10-x} = \binom{10}{x}\left(\frac{1}{2}\right)^{10}; x = 0, , 2, \ldots, 10$$

$\therefore$ Probability of getting at least seven heads is given by

$$P(X \geq 7) = p(7) + p(8) + p(9) + p(10)$$

$$= \left(\frac{1}{2}\right)^{10}\left\{\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right\}$$

$$= \frac{120 + 45 + 10 + 1}{1024} = \frac{176}{1024}.$$

Example 3

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that

(a) at least10 survive,

(b) from 3 to 8 survive, and

(c) exactly 5 survive?

Soln

(a) $$P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662$$

$$= 0.0338$$

(b) $$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4) = \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$$

$$= 0.9050 - 0.0271 = 0.8779$$

(c) $$P(X=5)=b(5;15,0.4)=\sum_{x=0}^{5} b(x;15,0.4)-\sum_{x=0}^{4} b(x;15,0.4)$$
$$=0.4032-0.2173=0.1859$$

Example 4

Out of 800 families with 4 children each, how many families would be expected to have

(i) 2 boys and 2 girls, (ii) at least 1 boy,

(iii) at most 2 girls and (iv) children of both sexes.

Assume equal probabilities for boys and girls.

Considering each child as a trial, n = 4. Assuming that birth of a boy is a success, p = ½ and q = 1/2

Let X denote the number of successes (boys).

Solution (i) $P(\text{2 boys and 2 girls}) = P(X=2)$

$$= 4C_2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{4-2}$$

$$= 6 \times \left(\frac{1}{2}\right)^4 = \frac{3}{8}$$

∴ No. of families having 2 boys and 2 girls

$= N \cdot (P(X=2)$ (where N is the total no. of families considered)

$$= 800 \times \frac{3}{8}$$

$= 300.$

(ii) $P(\text{at least 1 boy}) = P(X \geq 1)$

$$= P(X=1)+P(X=2)+P(X=3)+P(X=4)$$

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

$$= 1 - 4C_0 \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^4$$

$$= 1 - \frac{1}{16} = \frac{15}{16}$$

∴ No. of families having at least 1 boy

$$= 800 \times \frac{15}{16} = 750.$$

(iii) $P(\text{at most 2 girls}) = P(\text{exactly 0 girl, 1 girl or 2 girls})$

$$= P(X = 4, X = 3 \text{ or } X = 2)$$
$$= 1 - \{P(X = 0) + P(X = 1)\}$$
$$= 1 - \left\{ 4C_0 \cdot \left(\frac{1}{2}\right)^4 + 4C_1 \cdot \left(\frac{1}{2}\right)^4 \right\}$$
$$= \frac{11}{16}$$

$\therefore$ No. of families having at most 2 girls

$$= 800 \times \frac{11}{16} = 550.$$

Example 5

Example: For a binomial distribution, the mean is 6 and standard deviation is $\sqrt{2}$. Find the first two terms of distribution.

Solution: Mean $= 6 \Rightarrow np = 6$

Standard deviation $= \sqrt{2} \Rightarrow npq = 2$

By dividing, we get $q = \frac{1}{3}$

Hence, $p = \frac{2}{3}$

Therefore, $n = 9$

Required $= P(X = 0)$, $P(X = 1)$

$$P(X = 0) = {}^9C_0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^9 = \frac{1}{3^9}$$

$$P(X = 1) = {}^9C_1 \left(\frac{2}{3}\right)^1 \left(\frac{1}{3}\right)^8 = \frac{18}{3^9}$$

Example 6

Example: Comment on the statement: "The mean of a Binomial distribution is 5 and standard deviation is 3".

Solution:

Mean = 5

$\Rightarrow np = 5$

S.D = 3 $\Rightarrow Variance = 9$

$\Rightarrow npq = 9$

Divided these equations, we get

$$q = \frac{9}{5} > 1$$

which is not possible.

Hence, the given statement is not correct.

MOMENT GENERATING FUNCTION (m.g.f.)

The moment generating function (m.g.f.) of a random variable X is denoted by $M_X(t)$ and is defined as

$$M_X(t) = E(e^{tX}), \quad t \in R$$

wherever this expectation exists.

For discrete random variable

$$M_X(t) = \sum_x e^{tx} p(x)$$

Where $p(x)$ is the probability mass function (p.m.f.)

For any positive integer r, we denote $\mu_r = E(X^r)$

As its name implies, the m.g.f. can be used to compute a distribution's moments:

- the n^{th} moment about 0 is the n^{th} derivative of the moment-generating function, evaluated at 0.

i.e.

$$\mu_r = \frac{d^r}{dt^r}\left(M_X(t)\right) \quad at\ t = 0$$

Mean $= E(X)$

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

Example 7

Example: Find the m.g.f of the Binomial distribution and hence find its mean and Variance.

Solution: The p.m.f. for the Binomial distribution is

$$p(x) = {}^nC_x p^x q^{n-x}\ ; \quad x = 0,1,2,\dots,n;\ p + q = 1$$

Its m.g.f. is

$$M_X(t) = E(e^{tX})$$
$$= \sum_{x=0}^{n} e^{tx}\, {}^nC_x p^x q^{n-x}$$
$$= \sum_{x=0}^{n} {}^nC_x (e^t p)^x q^{n-x}$$
$$= (e^t p + q)^n$$

By using Binomial expansion of $(a + x)^n$,

provided $|x| < a$

Thus,

$$M_X(t) = (e^t p + q)^n$$

provided $|e^t p| < q$

For Mean and Variance $\quad M_X(t) = (e^t p + q)^n$

$$E(X) = \frac{d}{dt}(e^t p + q)^n$$
$$= n(e^t p + q)^{n-1} p e^t$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)$$
$$= \frac{d}{dt}\left[np\,(e^t p + q)^{n-1} e^t\right]$$
$$= np\begin{bmatrix}(n-1)(e^t p + q)^{n-2}(pe^t)e^t \\ +(e^t p + q)^{n-1} e^t\end{bmatrix}$$

At $t = 0$, we have

$$E(X) = n(p+q)^{n-1} p$$
$$= np$$
$$E(X^2) = np\begin{bmatrix}(n-1)(p+q)^{n-2}p \\ +(p+q)^{n-1}\end{bmatrix}$$
$$= np[(n-1)p + 1]$$

Thus,

Mean $= np$

Variance $= E(X^2) - [E(X)]^2$
$$= np[(n-1)p+1] - n^2p^2$$
$$= np(1-p)$$
$$= npq$$

Example 8

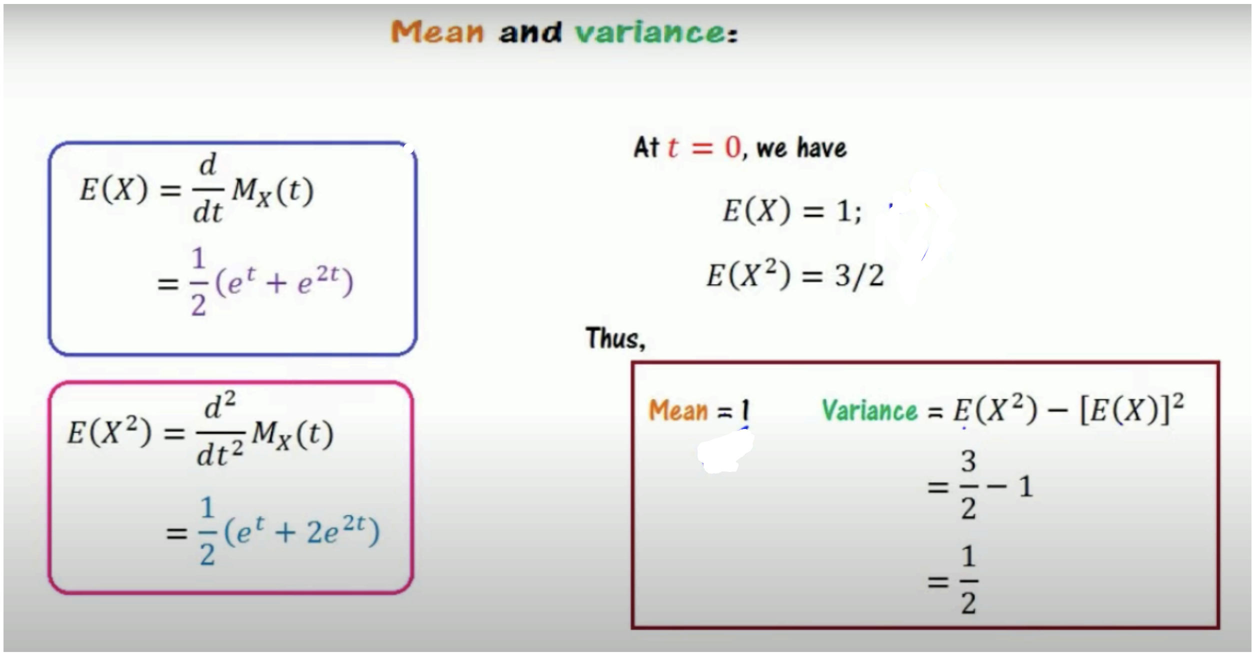
Example: A perfect coin is tossed twice. Find the m.g.f. of the number of heads. Hence, find the mean and variance.

Solution: The p.m.f. of X, (the number of heads) is

x	0	1	2
$p(x)$	¼	½	¼

The m.g.f. is

$$M_X(t) = E(e^{tX})$$
$$= \sum e^{tx} p(x)$$
$$= e^0\left(\frac{1}{4}\right) + e^t\left(\frac{1}{2}\right) + e^{2t}\left(\frac{1}{4}\right)$$
$$= \frac{1 + 2e^t + e^{2t}}{4}$$
$$= \frac{(1+e^t)^2}{4}$$



Example 9

Find the mean and variance of the binomial random variable of Example 3, and then use Chebyshev's theorem to interpret the interval $\mu \pm 2\sigma$.

Soln

$n = 15$ and $p = 0.4$

$$\mu = (15)(0.4) = 6 \text{ and } \sigma^2 = (15)(0.4)(0.6) = 3.6.$$

$\sigma = 1.897.$

the required interval $\mu \pm 2\sigma.$

$6 \pm (2)(1.897)$

By Chebyshev's theorem

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

Here k=2

So

$1-\frac{1}{k^2}$ = 1- ¼ = 3/4

P(2.206 < X< 9.794) ≥ ¾

i.e.

Chebyshev's theorem states that the number of recoveries among 15 patients who contracted the disease has a probability of at least 3/4 of falling between 2.206 and 9.794
or, because the data are discrete, between 2 and 10 inclusive.

Example 10

Example 7·2. *A and B play a game in which their chances of winning are in the ratio 3 : 2. Find A's chance of winning at least three games out of the five games played*

Solution. Let p be the probability that 'A' wins the game. Then we are given $p = 3/5 \Rightarrow q = 1 - p = 2/5$.

Hence, by binomial probability law, the probability that out of 5 games played, A wins 'r' games is given by :

$$P(X = r) = p(r) = \binom{5}{r} \cdot (3/5)^r\, (2/5)^{5-r}\,;\ r = 0, 1, 2, \ldots, 5$$

The required probability that 'A' wins at least three games is given by :

$$P(X \geq 3) = \sum_{r=3}^{5} \binom{5}{r} \frac{3^r \cdot 2^{5-r}}{5^5}$$

$$= \frac{3^3}{5^5}\left[\binom{5}{3} 2^2 + \binom{5}{4} . 3 \times 2 + 1.3^2 \times 1\right] = \frac{27 \times (40 + 30 + 9)}{3125} = 0{\cdot}68$$

Example 11

Example 7·8. *The mean and variance of binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $P(X \geq 1)$.*

Solution. Let $X \sim B(n, p)$. Then we are given

$$\text{Mean} = E(X) = np = 4$$

and $$\text{Var}(X) = npq = \tfrac{4}{3} \qquad \ldots(*)$$

Dividing, we get

$$q = \tfrac{1}{3} \quad \Rightarrow \quad p = \tfrac{2}{3}$$

Substituting in (*), we get

$$n = \frac{4}{p} = \frac{4 \times 3}{2} = 6.$$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - q^n = 1 - (1/3)^6 = 1 - (1/729)$$
$$= 1 - 0{\cdot}00137 = 0{\cdot}99863$$

Note

m.g.f. of Bernoulli variate (X=0, 1) is given by

$$M_X(t) = e^{0 \cdot t} \times P(X = 0) + e^{1 \cdot t} \cdot P(X = 1) = q + pe^t$$

For Bernoulli dist.
Mean= p
Variance=pq