3.4 Binomial Probability Distribution

Binomial Experiment:

An experiment for which the following 4 conditions are satisfied is called a binomial experiment.

1) The experiment consists of sequence of n smaller experiments called trials, where n is fixed in advance of the experiment.

2) Each trial can result in one of the same two possible outcomes, which we generally denote by success (S) and failure (F).

3) The trials are independent, so that the outcome on any particular trial does not influence the outcome on any other trial.

4) The probability of success $P(S)$ is constant from trial to trial, we denote this probability by p and probability of failure is denoted by $q (= 1-p)$.

Binomial random variable:

The binomial random variable X associated with a binomial experiment consisting of n trials is defined as

X = the number of success (S) among the n trials

Example:

For $n=3$, there are eight possible outcomes for the experiment.

$$\{SSS, SSF, SFF, FSS, FSF, FFS, SFS, FFF\}$$

Then from definition $X(SSS)=3$, $X(SSF)=2$, $X(SFF)=1$, ----, $X(FFF)=0$

$\therefore$ possible values for X in 3-trial experiment are $x=0,1,2,3$.

Similarly possible values for X in an n-trial experiment are $x=0,1,2,----,n$.

Some notations:

1) $p \to$ probability of success in a single trial
 $q=1-p \to$ probability of failure " " "

2) $X \sim Bin(n,p) \to$ X is a binomial r.v. based on n trials with probability of success p

3) $b(x;n,p) \to$ pmf of Binomial distribution with parameters n and p and r.v. X.
 $B(x;n,p) \to$ cdf of binomial distribution

4) $S \to$ Success, the occurrence of any event A
 $F \to$ failure, the non-occurrence of A

The Binomial Distribution:

A discrete random variable X is said to have a binomial distribution with parameters n (+ve integer) and p $(0<p<1)$ if pmf of X is given by

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

$$\binom{n}{x} = {}^nC_x = \frac{n!}{x!\,(n-x)!}$$

The cdf of binomial distribution:

If $X \sim Bin(n, p)$ then the cdf is given by

$$B(x; n, p) = P(X \le x) = \sum_{y=0}^{x} b(y; n, p), \quad x = 0, 1, 2, \cdots, n.$$

Theorem: (Mean and variance of Binomial distribution)
If $X \sim Bin(n, p)$, then $E(X) = np$, $V(X) = np(1-p) = npq$, and $\sigma_x = \sqrt{npq}$, where $q = 1-p$.

Proof: If $X \sim Bin(n, p)$ then the pmf is

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

$\therefore$ Mean, $\mu = E(X) = \sum_x x\, p(x)$, $p(x)$ = probability function

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$\Rightarrow \mu = E(X) = \sum_{x=1}^{n} \frac{n.(n-1)!}{(x-1)!(n-x)!} p^x q^{n-x}, \quad \text{if } x-1 = r$$

$$= \sum_{r=0}^{n-1} \frac{n(n-1)!}{r!(n-r-1)!} p^{r+1} q^{n-r-1}$$

$$= np \sum_{r=0}^{n-1} \binom{n-1}{r} p^r q^{n-1-r} = np(p+q)^{n-1} = np$$

$$\Rightarrow \boxed{E(X) = np} \quad [\because q = 1-p \Rightarrow p+q = 1, \; (a+b)^n = \sum_{r=0}^{n} {}^nC_r\, a^r b^{n-r}$$

To find $V(X)$, we knows that $V(X) = E\{X(X-1)\} - \mu(\mu-1)$, $\mu = E(X)$.

Now, $E\{X(X-1)\} = \sum_x x(x-1)\, p(x) = \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x q^{n-x}$

$$= \sum_{x=2}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x q^{n-x}, \quad \text{let } x-2 = r$$

$$= \sum_{r=0}^{n-2} \frac{n!}{r!(n-2-r)!} p^{r+2} q^{n-2-r}$$

$$\Rightarrow E\{x(x-1)\} = n(n-1)p^2 \sum_{r=0}^{n-2} \frac{(n-2)!}{r!(n-2-r)!} p^r q^{n-2-r}$$

$$= n(n-1)p^2 (p+q)^{n-2} = n(n-1)p^2 \quad [\because p+q=1]$$

$$\therefore V(x) = E\{x(x-1)\} - \mu(\mu-1), \quad \mu = E(x)$$

$$= n(n-1)p^2 - np(np-1), \quad [\because \mu = E(x) = np]$$

$$= \cancel{n^2p^2} - np^2 - \cancel{n^2p^2} + np$$

$$= np(1-p) = npq \quad [\because q = 1-p]$$

$$\therefore \sigma_x = \sqrt{V(x)} = \sqrt{npq} \quad (\text{standard deviation})$$