

Binomial Distributions

- It is also known as the "Bernoulli Distribution."
- Discovered by Swiss mathematician James Bernoulli in 1713 and was first published in 1713, eight years after his death.

The distribution can be used for the following conditions:

- (1) The random experiment is performed repeatedly a finite and fixed number of times.
- (2) The outcome of each trial can be classified into two mutually disjoint categories called \downarrow Failure and Success.
- (3) All the trials are independent.
- (4) The probability of success in any trial is P and is constant for each trial and $q = 1 - P$ is the probability of failure in any trial and is constant for each trial.

If X denotes the number of successes in n trials satisfying the above conditions, then X is a random variable which can take the values $0, 1, 2, \dots, n$.

$$P(r) = P(X=r) = {}^n C_r P^r q^{n-r}$$

To n trials, the total numbers of possible ways of obtaining r success and $(n-r)$ failure is

$$\frac{n!}{r!(n-r)!} = {}^n C_r$$

Formula of Binomial Distribution

r	$P(X=r) = P(r)$	Probabilities are the successive terms in the binomial expansion $(q+P)^n$ called the binomial distribution.
0	${}^n C_0 P^0 q^{n-0} = q^n$	
1	${}^n C_1 P^1 q^{n-1}$	
2	${}^n C_2 P^2 q^{n-2}$	
\vdots	\vdots	
n	${}^n C_n P^n q^{n-n} = P^n$	

$$\begin{aligned}
 \text{Mean} &= \sum r p(r) = {}^n C_1 P q^{n-1} + 2 {}^n C_2 q^{n-2} P^2 + 3 {}^n C_3 q^{n-3} P^3 + \dots + n P^n \\
 &= n q^{n-1} P + 2 \frac{n(n-1)}{2!} q^{n-2} P^2 + \frac{3n(n-1)(n-2)}{3!} q^{n-3} P^3 + \dots + n P^n \\
 &= np \left[q^{n-1} + (n-1) q^{n-2} P + \frac{(n-1)(n-2)}{2!} q^{n-3} P^2 + \dots + P^{n-1} \right] \\
 &= np \left[q^{n-1} + {}^{n-1} C_1 q^{n-2} P + {}^{n-1} C_2 q^{n-3} P^2 + \dots + P^{n-1} \right] \\
 &= np (q+P)^{n-1} \quad \begin{array}{l} \text{[By Binomial expansion for positive} \\ \text{integer index]} \end{array} \\
 &= np \quad [P+q=1]
 \end{aligned}$$

$$\begin{aligned}
 \text{Variance} &= \sum r^2 p(r) - [\sum r p(r)]^2 = \sum r^2 p(r) - (\text{mean})^2 \quad \text{--- (1)} \\
 \sum r^2 p(r) &= 1^2 \cdot {}^n C_1 q^{n-1} P + 2^2 {}^n C_2 q^{n-2} P^2 + 3^2 {}^n C_3 q^{n-3} P^3 + \dots + n^2 P^n \\
 &= n q^{n-1} P + \frac{4n(n-1)}{2} q^{n-2} P^2 + \frac{9n(n-1)(n-2)}{3!} q^{n-3} P^3 + \dots + n^2 P^n \\
 &= np \left[q^{n-1} + 2(n-1) q^{n-2} P + \frac{3}{2}(n-1)(n-2) q^{n-3} P^2 + \dots + np^{n-1} \right] \\
 &= np \left[\left\{ q^{n-1} + (n-1) q^{n-2} P + \frac{(n-1)(n-2)}{P^2} q^{n-3} P^2 + \dots + P^{n-1} \right\} + \right. \\
 &\quad \left. \{(n-1) q^{n-2} P + (n-1)(n-2) q^{n-3} P^2 + \dots + (n-1) P^{n-1}\} \right] \\
 &= np \left\{ \{(q+P)^{n-1}\} + (n-1) P \{ q^{n-2} + (n-2) q^{n-3} P + \dots + P^{n-2} \} \right\} \\
 &= np \left[(q+P)^{n-1} + (n-1) P (q+P)^{n-2} \right] \\
 &= np \left[(q+P)^{n-1} + (n-1) P (q+P)^{n-2} \right] \\
 &= np \left[1 + (n-1) P \right] \quad \left[\because P+q=1 \right]
 \end{aligned}$$

Substituting in (1) we get

$$\begin{aligned}
 \text{Variance} &= np \left[1 + np - P \right] - (np)^2 \\
 &= np \left[1 + np - P - np \right] \\
 &= np \left[1 - P \right] = npq \quad \left[\because 1 - P = q \right]
 \end{aligned}$$

As q is the probability of failure, we always have $0 \leq q \leq 1$.

$$\text{Variance} = np \times q < np = \text{Mean}$$

$$\text{Variance} < \text{Mean}$$

Reference: Fundamentals of Statistics by S.C. Gupta.