

BINOMIAL DISTRIBUTION

Let us consider “n” independent trials. If the successes (S) and failures (F) are recorded successively as the trials are repeated we get a result of the type

S S F F S ... F S

Let “x” be the number of success and hence we have (n – x) number of failures.

$$\begin{aligned}
 P(S S F F S \dots F S) &= P(S) P(S) P(F) P(F) P(S) \dots P(F) P(S) \\
 &= p p q q p \dots q p \\
 &= p p \dots p \times q q q \dots q \\
 &= x \text{ factor} \times (n - x) \text{ factors} \\
 &= p^x \cdot q^{n-x}
 \end{aligned}$$

But “x” success in “n” trials can occur in nC_x ways.

Therefore the probability of “x” successes in “n” trials is given by

$$P(X = x \text{ successes}) = nC_x p^x q^{n-x}, x = 0, 1, 2, \dots, n$$

Where $p + q = 1$

Assumptions in Binomial Distribution:

- (i) There are only two possible outcomes for each trial (success or failure)

- (ii) The probability of a success is the same for each trail.
- (iii) There are “n” trials where “n” is constant.
- (iv) The “n” trails are independent.

Mean and variance of a Binomial Distribution:

- (i) Mean(μ) = np
- (ii) Variance(σ^2) = npq

The variance of a Binomial Variable is always less than its mean.

$$\therefore npq < np.$$

Find the moment generating function of binomial distribution and hence find the mean and variance.

Sol: Binomial distribution is $p(x) = {}^nC_x p^x q^{n-x}$, $x = 0, 1, 2, \dots, n$

To find Mean and Variance:

$$\begin{aligned} M_X(t) &= E(e^{tx}) = \sum_{x=0}^n e^{tx} P(x) \\ &= \sum_{x=0}^n e^{tx} {}^nC_x p^x q^{n-x} \\ &= \sum_{x=0}^n {}^nC_x (pe^t)^x q^{n-x} \quad \because \sum_{x=0}^n {}^nC_x a^x b^{n-x} = (a + b)^n \end{aligned}$$

$$M_X(t) = (pe^t + q)^n$$

$$\begin{aligned} \text{Mean } E(X) &= \left[\frac{d}{dt} [M_*(t)] \right]_{t=0} = \left[\frac{d}{dt} [(pe^t + q)^n] \right]_{t=0} \\ &= [n(pe^t + q)^{n-1}(pe^t + 0)]_{t=0} \end{aligned}$$

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

$$E(X) = np$$

$$\begin{aligned} E(X^2) &= \left[\frac{d^2}{dt^2} [M_X(t)] \right]_{t=0} \\ &= \left[\frac{d}{dt} [n(pe^t + q)^{n-1}(pe^t)] \right]_{t=0} = np \left[\frac{d}{dt} [(pe^t + q)^{(n-1)} e^t] \right]_{t=0} \\ &= np[(pe^t + q)^{n-1} e^t + e^t(n-1)(pe^t + q)^{n-2} pe^t]_{t=0} \\ &= np[(p + q)^{n-1} + (n-1)(p + q)^{n-2} p] \\ &= np[1 + (n-1)p] = np[1 + np - p] \\ &= np[1 - p + np] = np[q + np] = npq + n^2 p^2 \end{aligned}$$

$$E(X^2) = (np)^2 + npq$$

$$\text{Variance} = E(X^2) - [E(X)]^2$$

$$= (np)^2 + npq - (np)^2 = n^2 p^2 - n^2 p^2 + npq$$

$$\text{Variance} = npq$$

Problems based on Binomial Distribution:

$$\text{Mean} = np$$

$$\text{Variance} = npq$$

1. Criticize the following statements “ The mean of a binomial distribution is 5 and the standard deviation is 3”

Solution:

$$\text{Given mean} = np = 5 \quad \dots (1)$$

$$\text{Standard deviation} = \sqrt{npq} = 3$$

$$\Rightarrow \text{Variance} = npq = 9 \quad \dots (2)$$

$$\frac{(2)}{(1)} \Rightarrow \frac{npq}{np} = \frac{9}{5} = 1.8 > 1$$

Which is impossible. Hence, the given statement is wrong.

2. If $M_X(t) = \frac{(2e^t+1)^4}{81}$, then find Mean and Variance.

Solution:

$$\text{Given } M_X(t) = \frac{(2e^t+1)^4}{81}$$

$$\Rightarrow M_X(t) = \left(\frac{2}{3}e^t + \frac{1}{3}\right)^4$$

Comparing with MGF of Binomial Distribution, $M_X(t) = (pe^t + q)^n$, we get

$$p = \frac{2}{3} \text{ and } q = \frac{1}{3}, n = 4$$

$$(i) \text{ Mean} = np = 4 \times \frac{2}{3} = \frac{8}{3}$$

$$(ii) \text{ Variance} = npq = \frac{8}{3} \times \frac{1}{3} = \frac{8}{9}$$

3. Six dice are thrown 729 times. How many times do you expect atleast 3 dice to show a five or six.

Solution:

Given $n = 6$ and $N = 729$

Probability of getting (5 or 6) $p = \frac{2}{6} = \frac{1}{3}$

and $q = 1 - \frac{1}{3} = \frac{2}{3}$

$$P(X = x) = {}^nC_x p^x q^{n-x}, x = 0, 1, 2, \dots, n$$

$$= {}^6C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{6-x}, x = 0, 1, 2, \dots, 6$$

$$P(\text{atleast 3 dice to show a five or six}) = P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[{}^6C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^{6-0} + {}^6C_1 \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^{6-1} + {}^6C_2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{6-2} \right]$$

$$= 1 - [0.0877 + 0.2634 + 0.3292]$$

$$= 1 - 0.6803$$

$$= 0.3197$$

Number of times expecting atleast 3 dice to show 5 or 6 $= 729 \times 0.3197$

$$= 233 \text{ times}$$

4. A machine manufacturing screw is known to produce 5% defective. In a random sample of 15 screws, what is the probability that there are (i) exactly 3 defectives (ii) not more than 3 defectives.

Solution:

Given $n = 15$

$$p = 5\% = 0.05$$

$$q = 1 - p = 1 - 0.05 = 0.95$$

$$\begin{aligned}
 P(X = x) &= {}^nC_x p^x q^{n-x}, x = 0, 1, 2, \dots, n \\
 &= {}^{15}C_x (0.05)^x (0.95)^{15-x}, x = 0, 1, 2, \dots, 15
 \end{aligned}$$

$$(i) P(\text{exactly 3 defectives}) = P(X = 3)$$

$$= {}^{15}C_3 (0.05)^3 (0.95)^{15-3}$$

$$= 0.056(0.95)^{12}$$

$$= 0.0307$$

$$(ii) P(\text{no More than 3 defectives}) = P(X \leq 3)$$

$$= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$= {}^{15}C_0 (0.05)^0 (0.95)^{15-0} + {}^{15}C_1 (0.05)^1 (0.95)^{15-1}$$

$$+ {}^{15}C_2 (0.05)^2 (0.95)^{15-2}$$

$$+ {}^{15}C_3 (0.05)^3 (0.95)^{15-3}$$

$$= {}^{15}C_0 (0.05)^0 (0.95)^{15} + {}^{15}C_1 (0.05)^1 (0.95)^{14} + {}^{15}C_2 (0.05)^2 (0.95)^{13}$$

$$+ {}^{15}C_3 (0.05)^3 (0.95)^{12}$$

$$= 0.4633 + 0.3658 + 0.1348 + 0.0307$$

$$= 0.9946$$

Poisson Distribution

Poisson Distribution is a limiting case of Binomial Distribution under the following assumptions.

- (i) The number of trials “n” should be independently large. i.e., $n \rightarrow \infty$
- (ii) The probability of successes “p” for each trail is indefinitely small.
- (iii) $np = \lambda$ should be finite where λ is a constant.

Application of Poisson Distribution:

Determining the number of calls received per minute at a call Centre or the number of unbaked cookies in a batch at a bakery, and much more.

Find the MGF for Poisson distribution and hence find the mean and variance.

Sol: Poisson distribution is $p(x) = \frac{e^{-\lambda} \lambda^x}{x!}$ $x = 0, 1, 2, \dots$,

$$M_X(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{-x} e^{-\lambda} \frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{(e^t)^x}{x!}$$

$$= e^{-\lambda} \left[1 + \frac{\lambda e^t}{1!} + \frac{(\lambda e^t)^2}{2!} + \frac{(\lambda e^t)^3}{3!} + \dots \right]$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{-\lambda + \lambda e^t} = e^{\lambda(e^t - 1)}$$

$$M_X(t) = e^{\lambda(e^t - 1)}$$

To find the mean and variance :

$$\text{Mean } E(X) = \left[\frac{d}{dt} [M_X(t)] \right]_{t=0}$$

$$= \left[\frac{d}{dt} [e^{\lambda(e^t - 1)}] \right]_{t=0} = [e^{\lambda(e^t - 1)} \lambda(e^t)]_{t=0} = e^{\lambda(e^0 - 1)} \lambda e^0 = e^0 \lambda$$

$$\text{Mean} = \lambda$$

$$\begin{aligned}
 E(X^2) &= \left[\frac{d^2}{dt^2} [M_X(t)] \right]_{t=0} = \left[\frac{d^2}{dt^2} [e^{\lambda(e^t-1)}] \right]_{t=0} = \left[\frac{d}{dt} [e^{\lambda(e^t-1)} \lambda e^t] \right]_{t=0} \\
 &= \lambda \left[\frac{d}{dt} [e^{\lambda(e^t-1)+t}] \right]_{t=0} \\
 &= \lambda [e^{\lambda(e^t-1)+t} (\lambda e^t + 1)]_{t=0} \\
 &= \lambda [e^0 (\lambda + 1)] \\
 &= \lambda (\lambda + 1) \\
 E(X^2) &= \lambda^2 + \lambda \\
 \text{Variance} &= E(X^2) - [E(X)]^2 \\
 &= \lambda^2 + \lambda - \lambda^2 \\
 \text{Variance} &= \lambda
 \end{aligned}$$

Problems based on Poisson Distribution:

1. Write down the probability mass function of the Poisson distribution which is approximately equivalent to B(100, 0.02)

Solution:

Given $n = 1000, p = 0.02$

$$\lambda = np = 1000 \times 0.02 = 20$$

The probability mass function of the Poisson distribution

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, \dots, \infty$$

$$= \frac{e^{-20} 20^x}{x!}; x = 0, 1, 2, \dots, \infty$$

2. One percent of jobs arriving at a computer system need to wait until weekends for scheduling, owing to core – size limitations. Find the probability that among a sample of 200 jobs there are no jobs that have to wait until weekends.

Solution:

Given $n = 200, p = 0.01$

$$\lambda = np = 200 \times 0.01 = 2$$

The Poisson distribution is

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, \dots, \infty$$

$P(\text{no jobs to wait until weekends}) = P(X = 0)$

$$P(X = 0) = \frac{e^{-2} 2^0}{0!} = e^{-2} = 0.1353$$

3. The proofs of a 500 pages book containing 500 misprints. Find the probability that there are atleast 4 misprints in a randomly chosen page.

Solution:

Given $n = 500$

$$p = P(\text{getting a misprint in a given page}) = \frac{1}{500}$$

$$\lambda = np = 500 \times \frac{1}{500} = 1$$

The Poisson distribution is

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, \dots, \infty$$

$$P(X \geq 4) = 1 - P(X < 4)$$

$$= 1 - [P(X = 0) + P(X = 1)] + P(X = 2) + P(X = 3)$$

$$= 1 - \left[\frac{e^{-1} 1^0}{0!} + \frac{e^{-1} 1^1}{1!} + \frac{e^{-1} 1^2}{2!} + \frac{e^{-1} 1^3}{3!} \right]$$

$$= 1 - e^{-1} \left[1 + 1 + \frac{1}{2} + \frac{1}{6} \right]$$

$$= 1 - 0.3679[2.666]$$

$$= 1 - 0.9809$$

$$= 0.0192$$