27.6 POISSON DISTRIBUTION

Poisson* distribution is the discrete probability distribution of a discrete random variable x, which has no upper bound. It is defined for non-negative values of x as follows:

$$f(x, \lambda) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for } x = 0, 1, 2, \ldots \quad (1)$$

Here $\lambda > 0$ is called the parameter of the distribution. Note that in binomial distribution the number of successes (occurrence of an event) out of a total definite number of n trials is determined, whereas in Poisson distribution the number of successes at a random point of time and space is determined.

Poisson distribution (P.D.) is suitable for 'rare' events for which the probability of occurrence p is very small and the number of trials n is very large. Also binomial distribution can be approximated by poisson distribution when $n \to \infty$ and $p \to 0$ such that $\lambda = np =$ constant.

Examples of rare events:

i. Number of printing mistakes per page.
ii. Number of accidents on a highway.
iii. Number of defectives in a production centre.
iv. Number of telephone calls during a particular (odd) time.
v. Number of bad (dishonoured) cheques at a bank.

Result 1: Since $\sum_{x=0}^{\infty} f(x,\lambda) = \sum_{x=0}^{\infty} p(X=x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$, Therefore (1) is a probability function.

Result 2: Arithmetic mean of Poisson distribution

$$\overline{X} = E(X) = \sum_{x=0}^{\infty} xP(X=x) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!}$$

$$= \lambda e^{-\lambda}\cdot e^{\lambda} = \lambda$$

Thus the parameter λ is the A.M. of P.D.

Result 3: Variance of Poisson distribution

$$= E\left[(x-\overline{X})^2\right] = \sum_{x=0}^{\infty}(x-\overline{X})^2 P(X=x)$$

$$= \sum\left(x^2 + \overline{X}^2 - 2\overline{X}x\right)P = \sum x^2 P + \overline{X}^2\sum P$$

$$-2\overline{X}\sum xP$$

$$= \sum x^2 P + \overline{X}^2 - 2\overline{X}\,\overline{X} = \sum x^2 P + \lambda^2 - 2\lambda^2$$

$$= \sum x^2 P - \lambda^2.$$

But

$$\sum x^2 P = \sum_{x=0}^{\infty}[x(x-1)+x]e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x\cdot x(x-1)}{x!} + e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x\cdot x}{x!}$$

$$= \lambda^2 e^{-\lambda}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!} + \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda^2 e^{-\lambda}e^{\lambda} + \lambda e^{-\lambda}e^{\lambda} = \lambda^2 + \lambda$$

Thus

$$\text{Variance} = \sum x^2 p - \lambda^2 = (\lambda^2+\lambda) - \lambda^2 = \lambda$$

Hence the variance of P.D. = mean of P.D.

Result 4: Recurrence formula

$$\frac{P(x+1)}{P(x)} = \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}\,\frac{x!}{e^{-\lambda}\lambda^x} = \frac{\lambda}{x+1}$$

Thus

$$P(x+1) = \left(\frac{\lambda}{x+1}\right)P(x).$$

Result 5: Poisson distribution function

$$F(x;\lambda) = \sum_{k=0}^{x}\frac{e^{-\lambda}\lambda^k}{k!}$$

has been tabulated (see A8 to A11)

Then $\quad f(x;\lambda) = F(x;\lambda) - F(x-1;\lambda).$

Theorem: *Prove that Poisson distribution is the limiting case of binomial distribution for very large trials with very small probability, i.e., $f(x;\lambda) = \lim_{\substack{n\to\infty \\ p\to 0}} b(x;n,p)$ such that $\lambda = np =$ constant.*

Proof: Put $p = \frac{\lambda}{n}$ in binomial distribution

$$b(x;n,p) = \frac{n!}{x!(n-x)!}\cdot\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)(n-2)\cdots(n-(x-1))}{x!}\cdot\frac{\lambda^x}{n^x}\times$$

$$\times\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n^x\cdot 1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{(x-1)}{n}\right)}{x!}\frac{\lambda^x}{n^x}\times$$

$$\times\left(1-\frac{\lambda}{n}\right)^{n-x}$$

Taking the limit as $n \to \infty$, we have

$$\lim_{n\to\infty} b(x;n,p) = \frac{1}{x!}\lambda^x \cdot e^{-\lambda} = \frac{\lambda^x e^{-\lambda}}{x!}$$

since

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right) = 1\cdot 1\cdots 1 = 1$$

and

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-x} = \left[\left(1-\frac{\lambda}{n}\right)^{n/\lambda}\right]^{\lambda} \times \left[1-\frac{\lambda}{n}\right]^{-x} = e^{-\lambda}.$$

Note: $\lim_{n\to\infty}\left(1-\frac{x}{n}\right)^n = e^{-x}$

Thus binomial probabilities for large n and small p are often approximated by means of poisson distribution with mean $\lambda = np$.

Example: For $n = 3000$, $p = 0.005$, the probability of 18 successes by binomial distribution is given by $b(18;3000,.005) = 3000_{C_{18}}(.005)^{18}(.995)^{2982}$ which involves prohibitive amount of work. Instead using Poisson distribution as an approximation, we get $\lambda = 3000 \times .005 = 15$. Probability of 18 success $= f(18, 15) = 0.8195$ from table (A8 to A11).

General rule: Poisson approximation to B.D. is used whenever $n \geq 20$ and $p \leq 0.05$. For $n \geq 100$, approximation is excellent provided $\lambda = np \leq 10$.

WORKED OUT EXAMPLES

Poisson distribution

Example 1: A distributor of bean seeds determines from extensive tests that 5% of large batch of seeds will not germinate. He sells the seeds in packets of 200 and guarantees 90% germination. Determine the probability that a particular packet will violate the guarantee.

Solution: The probability of a seed not germinating $= p = \frac{5}{100} = 0.05$

λ = mean number of seeds, in a sample of 200, which do not germinate

$= np = 200 \times 0.05 = 10$

Let X = R.V. = number of seeds that do not germinate

A packet will violate guarantee if it contains more than 20 non-germinating seeds.

Probability that the guarantee is violated

$$= P(X > 20) = 1 - P(X \leq 20) = 1 - \sum_{x=0}^{20}\frac{e^{-10}10^x}{x!}$$

$$= 1 - F(20, 10) = 1 - .9984 = 0.0016$$

where cummulative distribution function F is read for x=20 and λ=10 from the tables (A8 to A11).

Example 2: The average number of phone calls/minute coming into a switch board between 2 and 4 PM is 2.5. Determine the probability that during one particular minute there will be (a) 0 (b) 1 (c) 2 (d) 3 (e) 4 or fewer (f) more than 6 (g) at most 5 (h) at least 20 calls.

Solution: $\lambda = 2.5$, $f(x;\lambda) = f(x;2.5) = \frac{(2.5)^x(e^{-2.5})}{x!}$

Let X = R.V. = number of phone calls/minute during that (odd) 2 and 4 PM.

a. $f(0;2.5) = e^{-2.5} = .08208$

b. $f(1;2.5) = .2052$

c. $f(2;2.5) = .2565$

d. $f(3;2.5) = .2138$

e. $P(X \leq 4) = \sum_{x=0}^{4} f(x;2.5) = F(4;2.5) = .8912$

(read from tables A8 to A11)

f. $P(X > 6) = 1 - P(X \leq 6) = 1 - \sum_{x=0}^{6} f(x;2.5)$

$= 1 - F(6;2.5) = 1 - .9858 = 0.0142$

g. $P(X \leq 5) = \sum_{x=0}^{5} f(x;2.5) = F(5;2.5) = .9580$

h. $P(X \geq 2.0) = 1 - P(X \leq 19) = 1 - \sum_{x=0}^{19} f(x;2.5)$

$= 1 - F(19;2.5) = 1 - 1 = 0.$

Example 3: Suppose that on the average one person in 1000 makes a numerical error in preparing income tax return (ITR). If 10000 forms are selected at random and examined, find the probability that 6, 7 or 8 of the forms will be in error.

Example 26.43. *In a certain factory turning out razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10, use Poisson distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 10,000 packets.* (*Kurukshetra, 2009 S; Madras, 2006 ; V.T.U., 2004*)

Solution. We know that $m = np = 10 \times 0.002 = 0.02$

$$e^{-0.02} = 1 - 0.02 + \frac{(0.02)^2}{2!} - \ldots = 0.9802 \text{ approximately}$$

Probability of no defective blade is $e^{-m} = e^{-0.02} = 0.9802$

$\therefore$ no. of packets containing no defective blade is

$$10,000 \times 0.9802 = 9802$$

Similarly the number of packets containing one defective blade $= 10,000 \times me^{-m}$

$$= 10,000 \times (0.02) \times 0.9802 = 196$$

Finally the number of packets containing two defective blades

$$= 10,000 \times \frac{m^2 e^{-m}}{2!} = 10,000 \times \frac{(0.02)^2}{2!} \times 0.9802 = 2 \text{ approximately.}$$