

Lecture.7

Poisson Distributions - properties, Normal Distributions- properties

Theoretical Distributions

Theoretical distributions are

- | | | |
|--------------------------|---|-------------------------|
| 1. Binomial distribution | } | Discrete distribution |
| 2. Poisson distribution | | |
| 3. Normal distribution | → | Continuous distribution |

Discrete Probability distribution

Bernoulli distribution

A random variable x takes two values 0 and 1, with probabilities q and p ie., $p(x=1) = p$ and $p(x=0)=q$, $q=1-p$ is called a Bernoulli variate and is said to be Bernoulli distribution where p and q are probability of success and failure. It was given by Swiss mathematician James Bernoulli (1654-1705)

Example

- Tossing a coin(head or tail)
- Germination of seed(germinate or not)

Binomial distribution

Binomial distribution was discovered by James Bernoulli (1654-1705). Let a random experiment be performed repeatedly and the occurrence of an event in a trial be called as success and its non-occurrence is failure. Consider a set of n independent trials (n being finite), in which the probability p of success in any trial is constant for each trial. Then $q=1-p$ is the probability of failure in any trial.

The probability of x success and consequently n-x failures in n independent trials. But x successes in n trials can occur in nc_x ways. Probability for each of these ways is $p^x q^{n-x}$.

$$\begin{aligned} P(sss\dots ff\dots fsf\dots f) &= p(s)p(s)\dots p(f)p(f)\dots \\ &= p,p\dots q,q\dots \\ &= (p,p\dots p)(q,q\dots q) \\ &\quad (x \text{ times}) (n-x \text{ times}) \end{aligned}$$

Hence the probability of x success in n trials is given by

$$nc_x p^x q^{n-x}$$

Definition

A random variable x is said to follow binomial distribution if it assumes non-negative values and its probability mass function is given by

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

The two independent constants n and p in the distribution are known as the parameters of the distribution.

Condition for Binomial distribution

We get the binomial distribution under the following experimentation conditions

1. The number of trial n is finite
2. The trials are independent of each other.

3. The probability of success p is constant for each trial.
4. Each trial must result in a success or failure.
5. The events are discrete events.

Properties

1. If p and q are equal, the given binomial distribution will be symmetrical. If p and q are not equal, the distribution will be skewed distribution.
2. Mean = $E(x) = np$
3. Variance = $V(x) = npq$ (mean > variance)

Application

1. Quality control measures and sampling process in industries to classify items as defectives or non-defective.
2. Medical applications such as success or failure, cure or no-cure.

Example 1

Eight coins are tossed simultaneously. Find the probability of getting atleast six heads.

Solution

Here number of trials, $n = 8$, p denotes the probability of getting a head.

$$\therefore p = \frac{1}{2} \text{ and } q = \frac{1}{2}$$

If the random variable X denotes the number of heads, then the probability of a success in n trials is given by

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

$$= {}^8 C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{8-x} = {}^8 C_x \left(\frac{1}{2}\right)^8$$

$$= \frac{1}{2^8} {}^8 C_x$$

Probability of getting atleast six heads is given by

$$P(x \geq 6) = P(x = 6) + P(x = 7) + P(x = 8)$$

$$= \frac{1}{2^8} 8C_6 + \frac{1}{2^8} 8C_7 + \frac{1}{2^8} 8C_8$$

$$= \frac{1}{2^8} [8C_6 + 8C_7 + 8C_8]$$

$$= \frac{1}{2^8} [28 + 8 + 1] = \frac{37}{256}$$

Example 2 Ten coins are tossed simultaneously. Find the probability of getting (i) atleast seven heads (ii) exactly seven heads (iii) atleast seven heads

Solution

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

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

The probability of getting x heads throwing 10 coins simultaneously is given by

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

$$= {}^{10}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x} = {}^{10}C_x \left(\frac{1}{2}\right)^{10}$$

$$= \frac{1}{2^{10}} {}^{10}C_x$$

i) Probability of getting atleast seven heads

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

$$= \frac{1}{2^{10}} [{}^{10}C_7 + {}^{10}C_8 + {}^{10}C_9 + {}^{10}C_{10}]$$

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

ii) Probability of getting exactly 7 heads

$$P(x = 7) = \frac{1}{2^{10}} {}^{10}C_7 = \frac{1}{2^{10}} (120) = \frac{120}{1024}$$

iii) Probability of getting almost 7 heads

$$P(x \leq 7) = 1 - P(x > 7)$$

$$\begin{aligned}
&= 1 - \text{symbol } \{P(x=8) + P(x=9) + P(x=10)\} \\
&= 1 - \frac{1}{2^{10}} [10C_8 + 10C_9 + 10C_{10}] \\
&= 1 - \frac{1}{2^{10}} [45 + 10 + 1] \\
&= 1 - \frac{56}{1024} \\
&= \frac{968}{1024}
\end{aligned}$$

Example 3: 20 wrist watches in a box of 100 are defective. If 10 watches are selected at random, find the probability that (i) 10 are defective (ii) 10 are good (iii) at least one watch is defective (iv) at most 3 are defective.

Solution

20 out of 100 wrist watches are defective

$$\text{Probability of defective wrist watch, } p = \frac{20}{100} = \frac{1}{5}$$

$$q = 1 - p = \frac{4}{5}$$

Since 10 watches are selected at random, $n=10$

$$P(X=x) = {}^nC_x p^x q^{n-x}, x = 0, 1, 2, \dots, 10$$

$$= {}^{10}C_x \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{10-x}$$

i) Probability of selecting 10 defective watches

$$P(x=10) = {}^{10}C_{10} \left(\frac{1}{5}\right)^{10} \left(\frac{4}{5}\right)^0 = 1 \cdot \frac{1}{5^{10}} \cdot 1 = \frac{1}{5^{10}}$$

ii) Probability of selecting 10 good watches (i.e. no defective)

$$\begin{aligned}
P(x=0) &= {}^{10}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10} \\
&= 1 \cdot 1 \cdot \left(\frac{4}{5}\right)^{10} = \left(\frac{4}{5}\right)^{10}
\end{aligned}$$

iii) Probability of selecting at least one defective watch

$$P(x \geq 1) = 1 - P(x < 1)$$

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

$$= 1 - {}^{10}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10}$$

$$= 1 - \left(\frac{4}{5}\right)^{10}$$

iv) Probability of selecting at most 3 defective watches

$$P(x \leq 3) = P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)$$

$$= {}^{10}C_0 \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^{10} + {}^{10}C_1 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^9 + {}^{10}C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^8 + {}^{10}C_3 \left(\frac{1}{5}\right)^3 \left(\frac{4}{5}\right)^7$$

$$= 1.1 \left(\frac{4}{5}\right)^{10} + 10 \left(\frac{1}{5}\right)^1 \left(\frac{4}{5}\right)^9 + \frac{10 \cdot 9}{1 \cdot 2} \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^8 + \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} \left(\frac{1}{5}\right)^3 \left(\frac{4}{5}\right)^7$$

$$= 1. (0.107) + 10 (0.026) + 45 (0.0062) + 120 (0.0016)$$

$$= 0.859 \text{ (approx)}$$

Poisson distribution

The Poisson distribution, named after Simeon Denis Poisson (1781-1840). Poisson distribution is a discrete distribution. It describes random events that occurs rarely over a unit of time or space.

It differs from the binomial distribution in the sense that we count the number of success and number of failures, while in Poisson distribution, the average number of success in given unit of time or space.

Definition

The probability that exactly x events will occur in a given time is as follows

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

called as probability mass function of Poisson distribution.

where λ is the average number of occurrences per unit of time

$$\lambda = np$$

Condition for Poisson distribution

Poisson distribution is the limiting case of binomial distribution under the following assumptions.

1. The number of trials n should be indefinitely large i.e., $n \rightarrow \infty$
2. The probability of success p for each trial is indefinitely small.
3. $np = \lambda$, should be finite where λ is constant.

Properties

1. Poisson distribution is defined by single parameter λ .
2. Mean = λ
3. Variance = λ . Mean and Variance are equal.

Application

1. It is used in quality control statistics to count the number of defects of an item.
2. In biology, to count the number of bacteria.
3. In determining the number of deaths in a district in a given period, by rare disease.
4. The number of error per page in typed material.
5. The number of plants infected with a particular disease in a plot of field.
6. Number of weeds in particular species in different plots of a field.

Example 4: Suppose on an average 1 house in 1000 in a certain district has a fire during a year. If there are 2000 houses in that district, what is the probability that exactly 5 houses will have a fire during the year? [given that $e^{-2} = 0.13534$]

Solution:

Mean, $\bar{x} = np$, $n = 2000$ and $p = \frac{1}{1000}$

$$= 2000 \times \frac{1}{1000}$$

$$\lambda = 2$$

The Poisson distribution is

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\begin{aligned} P(X = 5) &= \frac{e^{-2} 2^5}{5!} \\ &= \frac{(0.13534) \times 32}{120} \\ &= \mathbf{0.036} \end{aligned}$$

Example 5

If 2% of electric bulbs manufactured by a certain company are defective. Find the probability that in a sample of 200 bulbs i) less than 2 bulbs ii) more than 3 bulbs are defective. [$e^{-4} = 0.0183$]

Solution

The probability of a defective bulb = $p = \frac{2}{100} = 0.02$

Given that $n = 200$ since p is small and n is large

We use the Poisson distribution

mean, $m = np = 200 \times 0.02 = 4$

Now, Poisson Probability function, $P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$

i) Probability of less than 2 bulbs are defective

$$\begin{aligned}
&= P(X < 2) \\
&= P(x = 0) + P(x = 1) \\
&= e^{-4} + e^{-4} (4) \\
&= e^{-4} (1 + 4) = 0.0183 \times 5 \\
&= 0.0915
\end{aligned}$$

ii) Probability of getting more than 3 defective bulbs

$$\begin{aligned}
P(x > 3) &= 1 - P(x \leq 3) \\
&= 1 - \{P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3)\} \\
&= 1 - e^{-4} \left\{ 1 + 4 + \frac{4^2}{2!} + \frac{4^3}{3!} \right\} \\
&= 1 - \{0.0183 \times (1 + 4 + 8 + 10.67)\} \\
&= 0.567
\end{aligned}$$

Normal distribution

Continuous Probability distribution is normal distribution. It is also known as error law or Normal law or Laplacian law or Gaussian distribution. Many of the sampling distribution like student-t, f distribution and χ^2 distribution.

Definition

A continuous random variable x is said to be a normal distribution with parameters μ and σ^2 , if the density function is given by the probability law

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

Note

The mean m and standard deviation s are called the parameters of Normal distribution. The normal distribution is expressed by $X \sim N(m, s^2)$

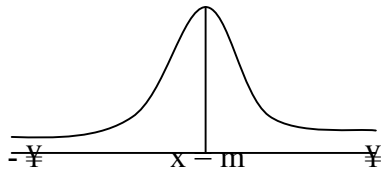
Condition of Normal Distribution

i) Normal distribution is a limiting form of the binomial distribution under the following conditions.

- a) n , the number of trials is indefinitely large i.e., $n \rightarrow \infty$ and
 - b) Neither p nor q is very small.
- ii) Normal distribution can also be obtained as a limiting form of Poisson distribution with parameter $m \rightarrow \infty$
- iii) Constants of normal distribution are mean = m , variation = s^2 , Standard deviation = s .

Normal probability curve

The curve representing the normal distribution is called the normal probability curve. The curve is symmetrical about the mean (m), bell-shaped and the two tails on the right and left sides of the mean extends to the infinity. The shape of the curve is shown in the following figure.



Properties of normal distribution

1. The normal curve is bell shaped and is symmetric at $x = m$.
2. Mean, median, and mode of the distribution are coincide
i.e., Mean = Median = Mode = m
3. It has only one mode at $x = m$ (i.e., unimodal)
4. The points of inflection are at $x = m \pm s$
5. The maximum ordinate occurs at $x = m$ and its value is $= \frac{1}{\sigma\sqrt{2\pi}}$
6. Area Property $P(m - s < x < m + s) = 0.6826$
 $P(m - 2s < x < m + 2s) = 0.9544$
 $P(m - 3s < x < m + 3s) = 0.9973$

Standard Normal distribution

Let X be random variable which follows normal distribution with mean μ and variance σ^2 . The standard normal variate is defined as $Z = \frac{X - \mu}{\sigma}$ which follows

standard normal distribution with mean 0 and standard deviation 1 i.e., $Z \sim N(0,1)$. The

standard normal distribution is given by $\phi(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$; $-\infty < z < \infty$

The advantage of the above function is that it doesn't contain any parameter. This enables us to compute the area under the normal probability curve.

Note

Property of $\phi(Z)$

1. $\phi(-Z) = 1 - \phi(Z)$
2. $P(a \leq Z \leq b) = \phi(b) - \phi(a)$

Example 6: In a normal distribution whose mean is 12 and standard deviation is 2. Find the probability for the interval from $x = 9.6$ to $x = 13.8$

Solution

Given that $Z \sim N(12, 4)$

$$\begin{aligned}
 P(9.6 \leq Z \leq 13.8) &= P\left(\frac{9.6 - 12}{2} \leq Z \leq \frac{13.8 - 12}{2}\right) \\
 &= P(-1.2 \leq Z \leq 0) + P(0 \leq Z \leq 0.9) \\
 &= P(0 \leq Z \leq 1.2) + P(0 \leq Z \leq 0.9) \quad [\text{by using symmetric property}] \\
 &= 0.3849 + 0.3159 \\
 &= 0.7008
 \end{aligned}$$

When it is converted to percentage (ie) 70% of the observations are covered between 9.6 to 13.8.

Example 7: For a normal distribution whose mean is 2 and standard deviation 3. Find the value of the variate such that the probability of the variate from the mean to the value is 0.4115

Solution:

Given that $Z \sim N(2, 9)$

To find X_1 :

We have $P(2 \leq Z \leq X_1) = 0.4115$

$$P\left(\frac{2-2}{3} \leq \frac{X-\mu}{\sigma} \leq \frac{X_1-2}{3}\right) = 0.4115$$

$$P(0 \leq Z \leq Z_1) = 0.4115 \text{ where } Z_1 = \frac{X_1 - 2}{3}$$

[From the normal table where 0.4115 lies is the value of Z_1]

From the normal table we have $Z_1 = 1.35$

$$\therefore 1.35 = \frac{X_1 - 2}{3}$$

$$\Rightarrow 3(1.35) + 2 = X_1$$

$$= X_1 = 6.05$$

(i.e) 41 % of the observation converged between 2 and 6.05

Questions

1. For a Poisson distribution

- | | |
|---------------------|---------------------|
| (a) mean > variance | (b) mean = variance |
| (c) mean < variance | (d) mean < variance |

Ans: mean = variance

2. In normal distribution, skewness is

- | | |
|----------------------|-------------------|
| (a) one | (b) zero |
| (c) greater than one | (d) less than one |

Ans: zero

3. Poisson distribution is a distribution for rare events

Ans: True

4. The total area under normal probability curve is one.

Ans: True

5. Poisson distribution is for continuous variable.

Ans: False

6. In a symmetrical curve mean, median and mode will coincide.

Ans: True

7. Give any two examples of Poisson distribution

8. The variance of a Poisson distribution is 0.5. Find $P(x = 3)$.

$[e^{-0.5} = 0.6065]$

9. The customer accounts of a certain departmental store have an average balance of Rs.1200 and a standard deviation of Rs.400. Assuming that the account balances are normally distributed. (i) what percentage of the accounts is over Rs.1500? (ii) What percentage of the accounts is between Rs.1000 and Rs.1500? iii) What percentage of the accounts is below Rs.1500?

10. State the Properties of normal distribution