Introduction to Probability Distribution

Random variable: A numerical value determined by the outcome of an experiment.
For example: Consider a random experiment in which a coin is tossed three times. Let *x* be the number of heads. Let *H* represent the outcome of a head and *T* the outcome of a tail. From the definition of a random variable, *x* as defined in this experiment is a *random variable*.

If we toss a coin, we may consider the random variable *x*, which is the number of heads 0 or 1.that is

Outcomes:	H	T
Random variables:	0	1

Random variable is of 2 types:

1. Discrete variable.
2. Continuous variable.

Probability Distribution: A listing of all possible outcomes of an experiment and the corresponding probability.
A probability distribution shows the possible outcomes of an experiment and the probability of each of these outcomes.

Example: Here's an example probability distribution that results from the rolling of a single fair die.

x	1	2	3	4	5	6	**sum**
p(x)	1/6	1/6	1/6	1/6	1/6	1/6	**6/6=1**

Types of Probability Distributions:
There are two types of probability distribution.
1. Discrete probability distribution
2. Continuous probability distribution

Discrete probability Distribution /probability mass function (pmf):
A function p(x) defined for a discrete random variable X is a probability mass function (pmf) if it satisfies the following 3 conditions-

i. $P(x)\geq 0$
ii. $\sum p(x)=1$
iii. $P(x)=P[X=x]$.

Discrete probability distribution can assume only certain outcomes. A probability distribution for a discrete a random variable is called a discrete probability distribution.
Example: The number of children in a family.

Properties of discrete distributions: Discrete distribution has some properties. Such as;

1. The sum of the probabilities of the various outcomes is 1.
2. The outcomes are mutually exclusive.
3. The probability of a particular outcome is between 0 and 1.

Some Discrete Distributions: There are many types of discrete probability distributions. Some of them are given below:

- Bernoulli Distribution
- Binomial Distribution
- Poisson Distribution

Binomial distribution

Introduction

Binomial distribution was first derived by Swiss mathematician James Bernoulli (1654-1705) and was first published posthumously in 1913, eight years after his death.

Definition

A discrete random variable X is said to have a binomial distribution if its probability function is defined by

$$f(x;n,p)=\begin{cases}\binom{n}{x}p^{x}q^{n-x} & \text{for } x=0,1,2,\ldots,n\\ 0; & \text{otherwise}\end{cases}$$

where the two parameters n and p satisfy $0\le p\le 1$ and $p+q=1$, also n is positive integer.

Mean: The mean of Binomial distribution, $\mu = np$.

Variance: The variance of Binomial distribution, $\sigma^2 = npq$.

Conditions for Binomial distribution

- Each trial results in two outcomes, termed as success and failure.
- The number of trials n is finite.
- The trials are independent of each other.
- The probability of success p is constant for each trial.

Example: A fair coin is tossed 5 times. Find the probability of (a) exactly two heads, (b) at least 4 heads,(c) at most 2 heard,(d)no heads, e) Find the mean and variance of that distribution.

Solution: Let the number of heads be random variate X which can take values 0, 1,2,3,4,and 5.Then X is a binomial variate with probability= $\frac{1}{2}$ and n=5.

Then the probability function of X is

$$f\left(x;5,\frac{1}{2}\right)=\binom{5}{x}\left(\frac{1}{2}\right)^{x}\left(\frac{1}{2}\right)^{5-x} \quad \text{for } x=0,1,2,3,4,5$$

(a) p[exactly two heads]=p[X=2]= $\binom{5}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{5-2}$

$=\frac{5!}{2!(5-2)!}(.25)(0.125)$

$=\frac{5\times4\times3!}{2!\times3!}(0.03125)$

=10*0.03125

=0.3125

(b) P [at least 4 heads] = $p[X\geq4]=p[X=4]+p[X=5]$

$=\binom{5}{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{5-4}+\binom{5}{5}\left(\frac{1}{2}\right)^{5}\left(\frac{1}{2}\right)^{5-5}$

=5*0.03125+0.03125

=0.15625+0.03125

=0.1875

(c) P [at most 2 heads]= $p[x\leq2]=p[x=2]+p[x=1]+p[x=0]$

$=\binom{5}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{3}+\binom{5}{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{4}+\binom{5}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{5}$

$=\frac{10}{32}+\frac{5}{32}+\frac{1}{32}=\frac{16}{32}=\frac{1}{2}$

(d) P [no heads] =p[X=0] = $\binom{5}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{5-0}=\frac{5!}{0!(5-0)!}\left(\frac{1}{2}\right)^{5}=\left(\frac{1}{2}\right)^{5}=0.03125$.

(e) We know,

Mean of binomial distribution is np
So
Mean is=np= $5\times\frac{1}{2}=2.5$

Variance is=npq=np(1-p)= $5\times\frac{1}{2}\times\frac{1}{2}=1.25$

Example

In a community, the probability that a newly born child will be boy $\frac{2}{5}$. Among the 4 newly born children in that community, what is the probability that

(a) All the four boys
(b) No boys
(c) Exactly one boy.

Solution

Let us consider the event that a newly born child is a boy as success in Bernoulli trial with probability of success $\frac{2}{5}$. Let the number of boys be a random variable X. Then X can take values 0, 1, 2, 3, and 4.

According to binomial law, the probability function of X is

$$f\left(x,4,\frac{2}{5}\right)=\binom{4}{x}\left(\frac{2}{5}\right)^{x}\left(\frac{3}{5}\right)^{4-x} \quad \text{for } x=0,1,2,3,4.$$

a) $p(\text{all boys})=p(x=4)=\binom{4}{4}\left(\frac{2}{5}\right)^{4}\left(\frac{3}{5}\right)^{4-4}=0.0256$.

b) $p(\text{no boys})=p(x=0)=\binom{4}{0}\left(\frac{2}{5}\right)^{0}\left(\frac{3}{5}\right)^{4-0}=0.1296$.

$p(\text{exactly one boy})=p(x=1)=\binom{4}{1}\left(\frac{2}{5}\right)^{1}\left(\frac{3}{5}\right)^{4-1}=0.3456$.

Example: The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of 6 workers
a) 4 or more will contract disease b) exactly 2 workers will contract disease?

Solution: The probability of a worker who is suffering from the disease i.e. $p = \frac{20}{100} = \frac{1}{5}$

The probability of a worker who is not suffering from the disease i.e.
$q = 1 - p = 1 - \frac{1}{5} = \frac{4}{5}$.

a) The probability of 4 or more i.e. 4, 5 or 6 will contract disease is given by

$$p[x \geq 4] = p[4] + p[5] + p[6]$$

$$= \binom{6}{4}\left(\frac{1}{5}\right)^4 \left(\frac{4}{5}\right)^{6-4} + \binom{6}{5}\left(\frac{1}{5}\right)^5 \left(\frac{4}{5}\right)^{6-5} + \binom{6}{6}\left(\frac{1}{5}\right)^6 \left(\frac{4}{5}\right)^{6-6}$$

$$= \binom{6}{4}\left(\frac{1}{5}\right)^4 \left(\frac{4}{5}\right)^2 + \binom{6}{5}\left(\frac{1}{5}\right)^5 \left(\frac{4}{5}\right) + \binom{6}{6}\left(\frac{1}{5}\right)^6$$

$$= 15\left(\frac{1}{5}\right)^4 \left(\frac{4}{5}\right)^2 + 6\left(\frac{1}{5}\right)^5 \left(\frac{4}{5}\right) + \left(\frac{1}{5}\right)^6$$

$$= 0.01696$$

b) The probability of exactly 2 workers will contract disease is given by

$$p[x = 2] = \binom{6}{2}\left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^{6-2}$$

$$= 15\left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^4$$

$$= 0.24576$$

Poisson distribution

Introduction

Poisson distribution was developed by France mathematician and physicist Simeon Denis Poisson (1781-1840), who published it in 1837.

Definition

A discrete random variable X is said to have a Poisson distribution if its probability function is given by

$$f(x;\lambda)=\begin{cases}\dfrac{e^{-\lambda}\lambda^{x}}{x!} & \text{for } x=0,1,2,\ldots,\infty.\\ 0; & \text{otherwise}\end{cases}$$

where, $e=2.71828$ and λ is the parameter of the distribution which is the mean number of success and $\lambda=np$.

Examples

- The number of cars passing a certain street in time t.
- Number of suicide reported in a particular day.
- Number of faulty blades in a packet of 100.
- Number of printing mistakes at each page of a book.
- Number of air accidents in some unit of time.
- Number of deaths from a disease such as heart attack or cancer or due to snake bite.
- Number of telephone calls received at a particular telephone exchange in some unit of time.
- The number of defective materials in a packing manufactured by a good concern.
- The number of letters lost in a mail on a given day in a certain big city.
- The number of fishes caught in a day in a certain city.
- The number of robbers caught on a given day in a certain city.

Example

Suppose that the number of emergency patients in a given day at a certain hospital is a Poisson variable X with parameter $\lambda=20$. What is the probability that in a given day there will be

a) 15 emergency patients.

b) At least 3 emergency patients.

c) More than 20 but less than 25 patients.

Solution

We know that,

$$f(x;\lambda)=\begin{cases}\dfrac{e^{-\lambda}\lambda^{x}}{x!} & \text{for } x=0,1,2,\ldots,\infty.\end{cases}$$

Here, $\lambda = 20$, $\therefore f(x;20) = \begin{cases} \dfrac{e^{-20}(20)^x}{x!} \end{cases}$ for $x = 0,1,2,...,\infty$.

a) $p(\text{15 emergency patients}) = p(x=15) = \dfrac{e^{-20}(20)^{15}}{15!} = 0.0516$.

b) $p(\text{at least 3 patients}) = p(x \geq 3) = 1 - p(x<3)$

$$= 1 - p(x=0) - p(x=1) - p(x=2)$$

$$= 1 - \frac{e^{-20}(20)^0}{0!} - \frac{e^{-20}(20)^1}{1!} - \frac{e^{-20}(20)^2}{2!} = 1.$$

c) $p(20 < x < 25) = p(x=21) + p(x=22) + p(x=23) + p(x=24)$

$$= \frac{e^{-20}(20)^{21}}{21!} + \frac{e^{-20}(20)^{22}}{22!} + \frac{e^{-20}(20)^{23}}{23!} + \frac{e^{-20}(20)^{24}}{24!} = 0.2841.$$

Example

If the probability that a car accident happens is a very busy road in on hour is 0.001. If 2000 cars passed in one hour by the road, what is the probability that

a) exactly 3
b) more than 2 car accidents happened on that hour of the road.

Solution

We know that,

$$f(x;\lambda) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} \end{cases} \text{ for } x = 0,1,2,...,\infty.$$

Here, $p = 0.001$, $n = 2000$. $\therefore \lambda = np = 2000 * 0.001 = 2$.

$$\therefore f(x;2) = \begin{cases} \dfrac{e^{-2}(2)^x}{x!} \end{cases} \text{ for } x = 0,1,2,...,\infty.$$

a) $p(\text{exactly 3 accidents}) = p(x=3) = \dfrac{e^{-2}(2)^3}{3!} = 0.18$.

b) $p(\text{more than 2 accidents}) = p(x>2) = 1 - p(x \leq 2)$

$$= 1 - p(x=0) - p(x=1) - p(x=2)$$

$$= 1 - \frac{e^{-2}(2)^0}{0!} - \frac{e^{-2}(2)^1}{1!} - \frac{e^{-2}(2)^2}{2!} = 0.323.$$

Example

A factory produces blades in a packet of 10. The probability of a blade to be defective is 0.2%. Find the number of packets having two defective blades in a consignment of 10,000 packets.

Solution

We know that,

$$f(x;\lambda) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & \text{for } x = 0,1,2,\ldots,\infty. \end{cases}$$

Here, $p = 0.2\% = 0.002$, $n = 10$. $\therefore \lambda = np = 10 * 0.002 = 0.02$.

$$\therefore p(\text{2 defective blades}) = p(x=2) = \frac{e^{-0.02}(0.02)^2}{2!} = 0.000196.$$

Therefore, the total number of packets having two defective blades in a consignment of 10,000 packet is $10000 \times 0.000196 = 1.96 \approx 2$.

Binomial vs. Poisson

Binomial Distribution	*Poisson Distribution*
Fixed Number of Trials (n) [10 pie throws]	Infinite Number of Trials
Only 2 Possible Outcomes [hit or miss]	Unlimited Number of Outcomes Possible
Probability of Success is Constant (p) [0.4 success rate]	Mean of the Distribution is the Same for All Intervals (mu)
Each Trial is Independent	Number of Occurrences in Any Given

[throw 1 has no effect on throw 2]	Interval Independent of Others