Random Variables

Random Variables

<u>Study Material for Week 6</u>

<u>Lecture Five</u>

- **Probability Models or Statistical Distributions**

In many practical situations, the random variable follows a specific pattern which can be described by a standard probability distribution. In these cases, probability mass function can be expressed in algebraic form and various characteristics of the distribution can be calculated using known closed formulae.

Bernoulli Distribution

Bernoulli random variableis an outcome of Bernoulli trial. Suppose, an experiment whose outcome can be classified as either a success or a failure. The probability of success is p and failure is $1-p$. Then the Bernoulli random variable X is defined as $X=1$ if trial is a success and $X=0$ if trial is a failure.

The distribution is discovered by Swiss Mathematician James Bernoulli, in 1713.

It is applied wherever the experiment results in only two outcomes.

One outcome is termed as a success 1 and the other as failure 0. Such an experiment is called <u>Bernoulli Trial</u>.

- **Probability Mass Function** $P(X=i)=\begin{cases} p^i(1-p)^{1-i} & i=0,1 \\ 0 & \text{otherwise} \end{cases}$

random variable $X=x_i$ 0 1 **probability mass function** $1-p$ p

- **Cumulative Distribution Function**

$$F(a)=\begin{cases} 0 & a<0 \\ 1-p & 0\le a<1 \\ 1 & 1\le a \end{cases}$$

- **Expectation / Mean** $E(X)=\sum_i i\,p(i)=0(1-p)+1\,p=p$

- **Variance** $Var(X)=E(X^2)-(E(X))^2=\sum_i i^2\,p(i)-p^2=p-p^2=p(1-p)$

- Applications

Situations where Bernoulli distribution is commonly used are

1) Items produced by a machine defective / non-defective.
2) Students appearing for examination pass / fail.

Binomial Distribution

Binomial random variable counts the number of successes when n Bernoulli trials are performed.Suppose n independent Bernoulli trials are performed. Each trial results in a success with probability p. If X represents the number of successes that occur in n trials, then X is said to be a Binomial random variable and the probability distribution is known as Binomial distribution with parameters $(n,\ p)$.

- **Probability Mass Function (p.m.f.)**

probability mass function of a binomial random variable having parameters $(n,\ p)$ is given by $P(X=i)={}^{n}C_{i}\,p^{i}(1-p)^{n-i},\quad i=0,1,\ldots,n$.

Note that $$\sum_{i=0}^{n}{}^{n}C_{i}\;p^{i}(1-p)^{n-i}=(p+1-p)^{n}=1.$$

- **Cumulative Distribution Function (c.d.f.)**

$$F(a)=P(X\le a)=\sum_{i=1}^{a}{}^{n}C_{i}\;p^{i}(1-p)^{n-i}$$

- **Expectation :** $E(X)=np$

$$E(X)=\sum_{i=0}^{n}i\;p(i)=\sum_{i=1}^{n}i\;{}^{n}C_{i}\;p^{i}\,(1-p)^{n-i}$$

$$i\;{}^{n}C_{i}=i\times\frac{n!}{i!(n-i)!}=\frac{n!}{(i-1)!(n-i)!}=\frac{n(n-1)!}{(i-1)!(n-i)!}=\frac{n(n-1)!}{(i-1)!((n-1)-(i-1))!}=n\;{}^{n-1}C_{i-1}$$

$$\therefore\;E(X)=\sum_{i=1}^{n}n\;{}^{n-1}C_{i-1}\;p^{i}\,(1-p)^{n-i}=n\sum_{i=1}^{n}{}^{n-1}C_{i-1}\;p^{i}\,(1-p)^{n-i},Put\;\,i-1=r$$

$$\therefore\;E(X)=n\sum_{r=0}^{n-1}{}^{n-1}C_{r}\;p^{r+1}\,(1-p)^{n-r-1}=np\sum_{r=0}^{n-1}{}^{n-1}C_{r}\;p^{r}\,(1-p)^{n-1-r}=np(p+1-p)^{n-1}=np$$

- **Variance**

$$E(X^{2})=\sum_{i=0}^{n}i^{2}\;p(i)=\sum_{i=0}^{n}i^{2}\;{}^{n}C_{i}\;p^{i}\,(1-p)^{n-i}=\sum_{i=1}^{n}i^{2}\;{}^{n}C_{i}\;p^{i}\,(1-p)^{n-i}$$

$$\sum_{r=0}^{n-1}r\;{}^{n-1}C_{r}\;p^{r}\,(1-p)^{n-1-r}=n\sum_{r=0}^{n-1}{}^{n-2}C_{r-1}\;p^{r}\,(1-p)^{n-1-r}$$

$$E(X^{2})=\sum_{i=1}^{n}(i^{2}-i+i)\,{}^{n}C_{i}\;p^{i}\,(1-p)^{n-i}=\sum_{i=1}^{n}i\,(i-1)\,{}^{n}C_{i}\;p^{i}\,(1-p)^{n+i}+\sum_{i=1}^{n}i\;{}^{n}C_{i}\;p^{i}\,(1-p)^{n-i}$$

Now,

$$i\,(i-1)\,{}^{n}C_{i}=\frac{i\,(i-1)\,n!}{i!(n-i)!}=\frac{n!}{(i-2)!(n-i)!}=n(n-1)\frac{(n-2)!}{(i-2)!((n-2)-(i-2))!}=n(n-1)\,{}^{n-2}C_{i-2}$$

$$\therefore\;\sum_{i=1}^{n}n(n-1)\,{}^{n-2}C_{i-2}\;p^{i}\,(1-p)^{n-i}=n(n-1)\sum_{i=2}^{n}{}^{n-2}C_{i-2}\;p^{i}\,(1-p)^{n-i}$$

$$i-2=r\quad\Rightarrow\quad n(n-1)\sum_{i=0}^{n-2}{}^{n-2}C_{r}\;p^{r+2}\,(1-p)^{n-r-2}=p^{2}n\,(n-1)\sum_{i=0}^{n-2}{}^{n-2}C_{r}\;p^{r}\,(1-p)^{(n-2)-r}$$

$= p^2 n(n-1)(p+1-p)^{n-2} = p^2 n(n-1)$

$\therefore E(X^2) = p^2 n(n-1) + np = p^2 n^2 - p^n n + np = p^2 n^2 + np(1-p)$

$\therefore Var(X) = E(X^2) - (E(X))^2 = \{p^2 n^2 + np(1-p)\} - n^2 p^2 = np(1-p) = npq$

Therefore $sd(X) = \sqrt{npq}$.

Mode of the distribution $= \lfloor (n+1)p \rfloor =$ integral part of $(n+1)p$.

- **Applications**

  Binomial distribution can be used when

  1. Trials are finite, performed repeatedly n times.
  2. Each trial should result as success or failure.
  3. Probability of success is constant for each trial as p .
  4. All trials are independent.

Example

1. The probability that a bomb dropped from a plane will strike the target is the probability that (i) exactly two will strike the target (ii) at least two will strike the target.

   Random Variable X : Count of bombs dropped from a plane striking the target

   $p = \frac{1}{5}$ & $n = 6 \therefore X \sim B\left(6, \frac{1}{5}\right)$.

   (i) Exactly two will strike the target.

   $$P(X = 2) = {}^6C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^4 = 15 \times (0.2)^6 = 15^3 \times \frac{4 \times 4 \times 4 \times 4}{5 \times 5 \times 5 \times 5 \times 5 \times 5} = 0.24576$$

   (ii) At least two will strike the target.

   $P(X \geq 2) = 1 - P(X < 2)$

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

   $= 1 - \left(\frac{1}{5}\right)^0 \left(\frac{4}{5}\right)^6 - 6 \times \left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^5$

   $= 1 - \left[\frac{4\times4\times4\times4\times4\times4}{5\times5\times5\times5\times5\times5} + \frac{6\times4^5}{5^6}\right]$

   $= 1 - \frac{4^5}{5^6}[4+6] = 1 - 10^2 \times \frac{4^5}{5\times5^5} = 1 - 0.65536 = 0.34464$

2. The probability that an entering student will graduate is 0.4. Determine the probability that out of 5 students (a) none (b) one (c) at least one will graduate.

   Random Variable X : Count of graduate students $p = 0.4,\ n = 5$

   $\therefore X \sim B(5, 0.4)$.

(a) $P(X=0)=(0.4)^0(0.6)^5=0.07776$

(b) $P(X=1)=0.2592$

(c) $P(X\geq 1)=1-P(X=0)=0.92224$

3. Point out the fallacy of the statement : The mean of binomial distribution is 3 and variance is 5.

$np=3$ & $npq=5 \quad q=\frac{5}{3}>1$

4. If the probability that a new born child is a male is 0.6, find the probability that in a family of 5 children there are exactly 3 boys.

X : count of a new born male child. Given $p=0.6$ & $n=5 \therefore X\sim B(5,0.6)$.

$$P(X=3)={}^5C_3(0.6)^3(1-0.6)^2=\frac{5!}{3!\,2!}(0.6)^3(0.4)^2=10\times(0.6)^3\times(0.4)^2=0.3456 \quad .$$

5. Out of 800 families with 5 children each, how many would you expect to have

i) 3 boys? ii) either 2 or 3 boys?

Assume equal probabilities for boys and girls.

Let X = Number of boys. Given $p=\frac{1}{2}$, $n=5$. $X\sim B\left(5,\frac{1}{2}\right)$.

i) $p(X=3)={}^5C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2=0.3125$

$\therefore$ Total number of families with 3 boys = $0.3125*800=250$

ii) $p(X=2\,or\,X=3)={}^5C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2+{}^5C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3=0.625$

$\therefore$ Total number of families with either 2 or 3 boys = $0.625*800=500$.

Poisson Distribution

A random variable X said to follow Poisson distribution if its probability mass function is given $p(X=r)=\frac{e^{-\lambda}\lambda^r}{r!}, r=0,1,2,3\cdots$. λ is known as parameter of the distribution.

- **Expectation and Variance**

If $X\sim P(\lambda)$ then $E(X)=\lambda, Var(X)=\lambda \therefore sd(X)=\sqrt{\lambda}$.

Poisson distribution as a limiting form of Binomial distribution

Poisson distribution can be derived as a limiting case of the Binomial distribution.

$X\sim B(n,p)$ if

i) The number of trials n is large, i.e., $n\to\infty$

ii) Probability of success is very small , i. e., $p \to 0$

iii) With this quantity $\lambda = np$ is finite.

- **Applications**

Poisson distribution can be used when

i) The rate of occurrence is constant over the time or space

ii) Past occurrences do not influence the likelihood of future occurrences.

iii) Simultaneous occurrences nearly impossible.

Following are few situations where Poisson distribution is applicable.

i) Arrival pattern of 'defective' items in a factory

ii) Arrival pattern of 'patients' in a hospital.

iii) Demand pattern of certain items

iv) Spatial distribution of bomb hits.

Examples

1. For a Poisson random variable X, $p(X=1) = 2P(X=2)$. Find the parameter λ of the distribution. Also find $p(X \geq 2)$.

$$p(X=1) = 2P(X=2) \Rightarrow \frac{e^{-\lambda}\lambda}{1!} = \frac{e^{-\lambda}\lambda^2}{2!} \Rightarrow \lambda = 2$$

$$p(X \geq 2) = 1 - p(X<2) = 1 - \left(p(X=0) + p(X=1)\right)$$

$$= 1 - \left(e^{-2} + 2e^{-2}\right) = 1 - 3e^{-2}$$

2. 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, two defective blades respectively in a consignment of 10,000 packets.

X : the blade to be defective. $p = 0.002,\ n = 10,\ \lambda = np = 10 \times 0.002 = 0.02$

Taking X as a Poisson random variable

i) $P(X=0) = \dfrac{e^{-0.02} \times (0.02)^0}{0!} = 0.980198673$

$\therefore$ No. of packets containing no defective blades

$$= 0.980198673 \times 10000 = 9801.98673$$

ii) $P(X=1) = \dfrac{e^{-0.02} \times (0.02)}{1} = 0.019603973$

iii) $P(X=2) = \dfrac{e^{-0.02} \times (0.02)^2}{2} = 1.960397346 \times 10^{-4}$

$\therefore$ No. of blades $= 1.960397346$

Random Variables

3. A manufacturer of cottor pins knows that 3% of his product is defective. If he sells cottor pins in boxes of 100 pins and guarantees that more than 5 pins will be defective in a box, find the approximate probability that a box will fail to meet the guaranteed quality.

X : product is defective. $p = 0.03,\ n = 100,\ \lambda = 3$

$P(X \geq 5) = 1 - P(X \leq 5)$

$$= 1 - \sum_{r=0}^{5} \frac{e^{-3} \times 3}{r!}$$

$$= 1 - 0.10081813 - 0.049787068 - 0.149361204$$

$$-0.02240418 - 0.224041806 - 0.340263493$$

$$= 0.113323436$$

Problem Session

Q. 1		**Attempt the following**
	1)	If the chance that one of the ten telephone lines is busy at an instant is 0.2 (a) What is the chance that 5 of the lines are busy? (b) What is the most probable number of busy lines and what is the probability of this number? **(c)** What is the probability that all the lines are busy?
	2)	A sortie of 20 aeroplanes is sent on an operational flight. The chances that an aeroplane fails to return is 5%. Find the probability that (i) one plane does not return (ii) at the most 5 planes do not return and (iii) what is the most probable number of returns?
Q. 2.		**Attempt the following**
	1)	If two probability of producing a defective screw is $P = 0.01$, what is the probability that a lot of 100 screws will contain more than 2 defectives?
	2)	A certain type of missile hits its target with probability $P = 0.3$. Find the number of missiles that should be fired so that there is at least a 90 percent probability of hitting the target.
	3)	A manufacturer knows that the condensers he makes contain on the average 1% defectives. He packs them in boxes of 100. What is the probability that a box picked at random will contain 3 or more faulty condensers?

Random Variables