SUBJECT:

PROBABILITY AND STATISTICS

For B.Tech II Year II Sem

Dr. K.RAMESH

Faculty of Mathematics
Department of Science & Humanities
SNIST

PROBABILITY AND STATISTICS

Unit-I

Random Variables and Probability Distributions

Dr. K.RAMESH

Faculty of Mathematics
Department of Science & Humanities
SNIST

- Conditional probability
- Multiplication theorem
- Baye’s Theorem (without Proof).
- Random variables –

  Discrete and Continuous random variables
- Probability Mass and Density functions
- Expectation and Variance.
- Probability Distributions:

  (i) Binomial Distribution,

  (ii) Poisson Distribution

  (iii) Normal Distribution.

Experiment: An operation which can produce some well defined outcome(s).(Or) Experiment is any physical action.

(i) Predictable experiment or Deterministic experiment (ii) Random experiment or Probabilistic experiment

Deterministic Experiment: When we perform an experiments in Science and Engineering and repeat the same under identical conditions we get the same result every time. This type of experiments called deterministic experiments.Here we can predict the result.

Random Experiment: when the experiment is repeated under the identical Conditions, if it do not produce the same outcome every time called Random experiment.

Examples: (1) Tossing a fair coin.

(2) Throwing a die (unbiased).

(3) Drawing a card from well shuffled pack of cards.

Sample space: The set of all possible outcomes in a random experiment is called Sample space. It is generally denoted by ‘S’.

Each element of sample space is called a sample point.

Examples: (1) In tossing a fair coin (R.E.), there are two possible outcomes head (H) and tail (T).
so the sample space is S = {H,T}   n(S)=2

Trail :Single performance of an experiment is called trail.

Events: The results or out comes of an experiment are called events.

*Events are generally denoted by the letter ‘E’ (Capital letter)

* If number of events are more we represent by $E_1, E_2, ...E_n$

Example: (1) Throwing an unbiased die (R.E.),Sample space S={1,2,3,4,5,6}

E1= getting even number ={2,4,6}

E2= getting odd number ={1,3,5}

Types of events: Events are two types 1.Elemantary events or simple events, 2.Compound events.

1.Elemantary event or simple event: Suppose that we have conducted a random experiment. It is completely defined when we know all possible out comes. Each out come of the experiment is called Elementary event or simple event.

Examples: (1) Getting head (H) and tail (T) are Elementary events when tossing a fair coin (R.E.),

2.Compound event: Joint occurrence of two or more elementary events is called Compound event.

Examples: Getting Even number when a fair die is rolled is Compound event

Impossible event: Let S be the sample space. Since ø (null set) is subset of S.

so 'ø' is called an impossible event.

Example: In throwing an unbiased die (R.E)

E1 = getting the positive integer more than 7 = ø
E2 = getting the positive integer less than 1 = ø

Sure event: Let S be the sample space. Since S is subset of S. so the event S

is called an impossible event.

Exhaustive events: Let S be a sample space of a random experiment and E_1 ,E_2, …..E_n be the events defined on S . If … then the events E_1 ,E_2,…..E_n are exhaustive events.

Equally likely events: The given events are said to be equally likely if none of them is expected to occur in preference to other.

Example: In tossing a coin (R.E)

E=getting head = {H} , F= getting tail= {T}

E and F are equally likely events.

Mutually Exclusive : Two events E and F are said to be mutually exclusive and exhaustive events if $E \bigcap F = \phi$

Example: Throwing an unbiased die (R.E.)

Sample space S={1,2,3,4,5,6}

E= getting even number ={2,4,6}

F= getting odd number ={1,3,5},

$E \bigcap F = \phi$ and $E \bigcup F = S$

So E and F are mutually exclusive and exhaustive events.

G= getting odd prime number = {3,5}

E and G are mutually exclusive but not exhaustive events.

Independent events: Two events are said to be independent if the occurrence of one does not depends on the occurrence of the other.

Example: Suppose two coins are tossed

E = event of occurrence of head on first coin

F= event of occurrence of Tail on second coin

Clearly the occurrence of head on second coin does not depends upon the occurrence of head on first coin.

So E and F are independent events.

* The events which are not independent are know as dependent events.

Mutually Exclusive and Exhaustive events: Two events E and F are said to be mutually exclusive and exhaustive events if $E \cap F = \phi$ and $E \cup F = S$

Example: Throwing an unbiased die (R.E.)

Sample space S={1,2,3,4,5,6}

E= getting even number ={2,4,6}

F= getting odd number ={1,3,5},

$E \cap F = \phi$ and $E \cup F = S$

So E and F are mutually exclusive and exhaustive events.

G= getting odd prime number = {3,5}

E and G are mutually exclusive but not exhaustive events.

Independent events: Two events are said to be independent if the occurrence of one does not depends on the occurrence of the other.

Example: Suppose two coins are tossed

E = event of occurrence of head on first coin

F= event of occurrence of Tail on second coin

Clearly the occurrence of head on second coin does not depends upon the occurrence of head on first coin.

So E and F are independent events.

* The events which are not independent are know as dependent events.

Random Variables

Prepared by

Dr.K.RAMESH

Random variable.

Let S be the samples space of some random experiment in which the elements need not be numbers. Now assigning the real values to the outcomes of a random experiment called *random variable*. Random variable is a single valued function.

Definition: Random variable is a mapping from sample space S to real number set R.

i.e. $X : S \longrightarrow R$ is a mapping then 'X' is a random variable.

Types of Random variable

Random variable are two types.

(i)Discrete Random variable
(ii)Continuous Random variable.

Discrete Random variable:

A random variable which can take only finite number of values or countably infinite number of values is called discrete random variable.

(OR)

A Random variable ‘X’ is said to be discrete random variable if its set of possible outcomes i.e the sample space, is countable.(finite or an unending sequence with as many elements as there are whole numbers or(countable)).

Examples of discrete random variables

- The number of children in a family.
- The number of patients in a doctor's surgery.
- The number of defective light bulbs in a box of ten.

Discrete Probability Distribution: Each element in a sample space has certain probability of occurrence. A formula representing all these probabilities which a discrete random variable assumes is known as the Discrete Probability distribution.

Suppose X is a discrete random variable with possible outcomes

(values) $x_1, x_2, x_3, \ldots$. The probability of each possible outcome x_i is

$$p_i = P(X = x_i) = P(x_i) \text{ for } i = 1, 2, 3, \ldots$$

The probability distribution table is given by

X	x_1	x_2	x_3	....	x_i	....	x_n
$P(X)$	p_1	p_2	p_3	....	p_i	....	p_n

Note:

(i) $P(X < x_i) = p(x_1) + p(x_2) + \dots + p(x_{i-1})$

(ii) $P(X \le x_i) = p(x_1) + p(x_2) + \dots + p(x_{i-1}) + p(x_i)$

(iii) $P(X > x_i) = 1 - P(X \le x_i)$

Example: Tossing two coins simultaneously (R.E)

X = Number of heads occurring (random variable).

The sample space, S = { HH, HT,TH,TT }

Therefore 'X' has the values 0,1,2.

i.e. Occurring of zero head, one head, two heads

The probabilities at various values of X are

$P(X=0) = P(\text{ zero heads}) = \frac{1}{4}$

$P(X=1) = P(\text{ one head}) = \frac{2}{4} = \frac{1}{2}$

$P(X=2) = P(\text{ two heads}) = \frac{1}{4}$

X= xi	0	1	2
P(X=xi)	1/4	1/2	1/4

The above table is the probability distribution table.

Probability Mass Function: Probability mass function of A discrete random variable 'X' is a function 'f(x)' [(Or) $P(X=x_i)=P(x_i)=p_i$)] satisfying the following conditions.

$$(i)\ f(x) \geq 0, \qquad i.e.\ P(X=x_i) \geq 0$$

$$(ii) \sum f(x) = 1 \qquad i.e. \sum P(X=x_i) = 1$$

$$i = 1, 2, 3, \ldots$$

Cumulative Distribution Function:

Cumulative distribution function(or **Distribution Function)** of a discrete random variable X is denoted by F(x) and is defined by

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

$$\text{for } -\infty < x < \infty.$$

It follows that

(i) $F(-\infty) = 0$

(ii) $F(\infty) = 1$

(iii) $P(x_i) = P(X = x_i) = F(x_i) - F(x_{i-1})$

Properties of Distribution Function :

1. If F is the distribution function of a random variable X and if $a < b$, then

(i) $P(a < X \leq b) = F(b) - F(a)$

(ii) $P(a \leq X \leq b) = P(X = a) + [F(b) - F(a)]$

(iii) $P(a < X < b) = [F(b) - F(a)] - P(X = b)$

(iv) $P(a \leq X < b) = [F(b) - F[(a)] - P(X = b) + P(X = a)$

Note : If $P(X = a) = P(X = b) = 0$ then

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = F(b) - F(a)$$

Expectation of Discreate Random variable:

Suppose a random variable X assumes the values $x_1, x_2, \ldots, x_n$ with respective probabilities $p_1, p_2, \ldots, p_n$. Then the **Mathematical Expectation** or **Mean** or **Expected value** of X, denoted by $E(X)$, is defined as the sum of products of different values of x and the corresponding probabilities.

$$\therefore\ E(X) = x_1 p_1 + x_2 p_2 + \ldots. + x_n p_n$$

$$\textit{i.e.,}\ \ E(X) = \sum_{i=1}^{n} p_i\, x_i$$

Variance of Discreate Random variable:

If X is a random variable, then the mathematical expectation of $(X-\mu)^2$ is defined to be the variance of the random variable X. Then

$$\text{Var}(X) = E(x-\mu)^2 = \sum_{i=1}^{n} p_i (x_i - \mu)^2$$

$$= \Sigma (x_i^2 - 2\mu x_i + \mu^2) p_i$$

$$= \Sigma x_i^2 p_i - 2\mu \Sigma x_i p_i + \mu^2 \Sigma p_i$$

$$= E(X^2) - 2\mu . \mu + \mu^2 .1 \quad [\because \mu = \Sigma x_i p_i, \Sigma p_i = 1]$$

$$\therefore \sigma^2 = E(X^2) - \mu^2 \quad i.e., \ \textbf{Var (X)} = \boldsymbol{E(X^2) - [E(X)]^2}$$

i.e $$Var(X) = \sigma^2 = \sum_{i=1}^{n} x_i^2 p_i - \mu^2$$

Standard Deviation : It is the positive square root of the variance.

$$\therefore \ \text{S.D} = \sigma = \sqrt{\sum_{i=1}^{n} p_i x_i^2 - \mu^2} = \sqrt{E(X^2) - \mu^2} = \sqrt{E[X - E(X)]^2}$$

1. A random variable X has the following probability function find (i) the value of K (ii) Mean (iii) $P(0<X<5)$ (iv) Cumulative distribution function(or **Distribution Function)**

X	0	1	2	3	4	5	6	7
P(X)	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Solution:

(i)We know that Sum of all probabilities must be equal to 1

For discreate random variable

$$\sum f(x)=1 \text{ or } \sum P(X = x_i)=p_i=1$$

So we get $K+2K+2K+3K+K^2+2K^2+7K^2+K=1$

$\Rightarrow 10K^2+9K=1$
$\Rightarrow K=-1, 1/10$

But probability cannot be negative , therefore the value of $K=1/10$

(ii)The mean
$\mu=1.K+2.2K+3.2K+4.3K+5K^2+6.2K^2+7.7K^2+7K$
$=30K+66K^2=3+0.66=3.66$

(iii)$P(0<x<5)=P(x=1)+P(x=2)+P(x=3)+P(x=4)=K+2K+2K+3K=8K=0.8$

(*iv*) The distribution function of X is given by the following table :

X	$F(x) = P(X \leq x)$
0	0
1	$K = 1/10$
2	$3K = 3/10$
3	$5K = 5/10$
4	$8K = 8/10$
5	$8K + K^2 = 81/100$
6	$8K + 3K^2 = 83/100$
7	$9K + 10K^2 = 1$

2.A random variable X has the following probability function (i) find the value of K (ii) Mean (iii) Variance

X	0	1	2	3	4	5	6
$P(X=x_i)P(X)=p_i$	0	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Solution : (*i*) Since the total probability is unity, we have $\sum_{x=0}^{6} p(x) = 1$

i.e., $0+2K+2K+3K+K^2+2K^2+(7K^2+K)=1$

i.e., $10K^2+8K-1=0$

$$\therefore \; K = \frac{-8 \pm \sqrt{64+40}}{20} = \frac{-8 \pm \sqrt{104}}{20} = \frac{-8 \pm 2\sqrt{26}}{20} = \frac{-4 \pm \sqrt{26}}{10}$$

Since, $p(x) \geq 0$,

$$\therefore \; K = \frac{-4+\sqrt{26}}{10} = 0.1099$$

(*ii*) Mean, $\mu = \sum_{i=0}^{6} p_i x_i$

$$= (0)(0) + (1)(2K) + (2)(2K) + (3)(3K) + (4)(K^2) + (5)(2K^2) + (6)(7K^2 + K)$$

$$= 2K + 4K + 9K + 4K^2 + 10K^2 + 42K^2 + 6K$$

$$= 56K^2 + 21K = K(56K + 21)$$

$$= (0.1099)[56(0.1099) + 21] = 2.9842$$

(*iii*) Variance $= \sum_{i=0}^{6} p_i x_i^2 - \mu^2$

$$= 0 + 2K(1)^2 + 2K(2)^2 + 3K(3)^2 + K^2(4)^2 + 2K^2(5)^2 + (7K^2 + K)(6)^2 - (2.9842)^2$$

$$= 2K + 8K + 27K + 16K^2 + 50K^2 + 252K^2 + 36K - 8.9054$$

$$= 318K^2 + 73K - 8.9054 = K(318K + 73) - 8.9054$$

$$= (0.1099)[318(0.1099) + 73] - 8.9054 = 2.9581$$

Example

Let X denotes the minimum of the two numbers that appear when a pair of fair dice is thrown once. Determine (i) Discrete probability distribution (ii) Expectation (iii) Variance.

Solution:

When two dice are thrown, total number of outcomes is 6x6=36

$$\text{In this case, sample space } S=\left\{\begin{matrix}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)\\(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\\(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\\(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\\(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)\\(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\end{matrix}\right\}$$

If the random variable X assigns the minimum of its number in S, then

$S^* = \{$ 1, 1, 1, 1, 1, 1,
1,2, 2, 2, 2, 2,
1, 2, 3, 3,3 ,3,
1,2,3, 4, 4,4
1,2,3, 4,5,5
1,2,3,4,5,6}

The minimum number could be 1,2,3,4,5,6.

For minimum 1, favourable cases are (1, 1), (1, 2), (2, 1), (1, 3), (3, 1), (1, 4), (4, 1), (1, 5), (5, 1), (1, 6), (6, 1).

So $P(X=1)=\frac{11}{36}$

For minimum 2, favourable cases are (2, 2), (2, 3), (3, 2), (2, 4), (4, 2), (2, 5) , (5, 2), (2, 6), (6, 2).

So $P(X = 2) = \frac{9}{36}$

Similarly, $P(X = 3) = P((3, 3), (3, 4), (4, 3), (3, 5), (5, 3), (3, 6), (6, 3)) = \frac{7}{36}$

$P(X = 4) = P((4, 4), (4, 5), (5, 4), (4, 6), (6, 4)) = \frac{5}{36}$

$P(X = 5) = P((5, 5), (5, 6), (6, 5)) = \frac{3}{36}$

$P(X = 6) = \mathrm{P}((6, 6)) = \frac{1}{36}$

$\therefore$ The probability distribution is

X	1	2	3	4	5	6
$P(X)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

(*ii*) Expectation = Mean = $\Sigma\, p_i x_i$

$$i.e.,\ E(X) = 1.\frac{11}{36} + 2.\ \frac{9}{36} + 3.\frac{7}{36} + 4.\frac{5}{36} + 5.\frac{3}{36} + 6.\frac{1}{36}$$

$$\text{or} \qquad \mu = \frac{1}{36}\,(11 + 18 + 21 + 20 + 15 + 6) = \frac{91}{36} = 2.5278$$

(iii) Variance $= \Sigma p_i x_i^2 - \mu^2$

$$= \frac{11}{36}.1 + \frac{9}{36}.4 + \frac{7}{36}.9 + \frac{5}{36}.16 + \frac{3}{36}.25 + \frac{1}{36}.36 - \left(\frac{91}{36}\right)^2$$

$$= \frac{1}{36}(11 + 36 + 63 + 80 + 75 + 36) - \left(\frac{91}{36}\right)^2$$

i.e., $\sigma^2 = 8.\ 3611 - 6.3898 = 1.9713$

Note : Standard deviation, $\sigma = \sqrt{1.9713} = 1.404$

Example:

A sample of 4 items are selected randomly from a box containing 12 items of which 5 are defective find the expected number of defective items.

<u>**Solution**</u> : Let X denote the number of defective items among 4 items drawn from 12 items.

Obviously X can take the values 0, 1, 2, 3, or 4.

No. of good items = 7

No. of defective items = 5

$$P(X=0) = P(\text{no defective}) = \frac{{}^{7}C_4}{{}^{12}C_4} = \frac{35}{495} = \frac{7}{99}$$

$P(X = 1) =$ P(one defective and 3 good items)

$$= \frac{{}^{7}C_3 \times {}^{5}C_1}{{}^{12}C_4} = \frac{7\times6\times5\times5}{6} = \frac{175}{495} = \frac{35}{99}$$

$P(X = 2) = P$(2 defective and 2 good items)

$$= \frac{{}^{7}C_2 \times {}^{5}C_2}{{}^{12}C_4} = \frac{210}{495} = \frac{42}{99}$$

$$P(X = 3) = P(3 \text{ defective and } 1 \text{ good item})$$

$$= \frac{{}^{7}C_1 \times {}^{5}C_3}{{}^{12}C_4} = \frac{70}{495} = \frac{14}{99}$$

$$P(X = 4) = P(\text{all are defective})$$

$$= \frac{{}^{5}C_4}{{}^{12}C_4} = \frac{5}{495} = \frac{1}{99}$$

Discrete probability distribution is

$X = x_i$	0	1	2	3	4
$P(X = x_i) = f(x_i)$	$\frac{7}{99}$	$\frac{35}{99}$	$\frac{42}{99}$	$\frac{14}{99}$	$\frac{1}{99}$

Expected number of defective items = $E(X) = \Sigma\, x_i . f(x_i)$

$$= 0.\frac{7}{99} + 1.\frac{35}{99} + 2.\frac{42}{99} + 3.\frac{14}{99} + 4.\frac{1}{99} = \frac{165}{99}$$

Example:

A discrete random variable X has the following distribution function :

$$F(x) = \begin{cases} 0, \ for\ x < 1 \\ 1/3, \ for\ 1 \le x < 4 \\ 1/2, \ for\ 4 \le x < 6 \\ 5/6, \ for\ 6 \le x < 10 \\ 1, \ for\ x \ge 10 \end{cases}$$

Find (i) $P(2 < X \le 6)$ (ii) $P(X = 5)$

(iii) $P(X \le 6)$ (iv) $P(X = 6)$

Solution :

(*i*) $P(2 < X \leq 6) = F(6) - F(2) = P(X \leq 6) - P(X \leq 2) = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$

(*ii*) $P(X = 5) = P(X \leq 5) - P(X < 5) = F(5) - P(X < 5) = \frac{1}{2} - \frac{1}{2} = 0$

(*iii*) $P(X \leq 6) = F(6) = \frac{5}{6}$

(*iv*) $P(X = 6) = F(6) - P(X < 6) = \frac{5}{6} - \frac{1}{2} = \frac{2}{6} = \frac{1}{3}$

Note: (1) In a gambling game, expected value E(X) of the game is considered to be the value of the game to the player.

The game is favourable to the player if $E>0$,
The game is unfavourable to the player if $E<0$,
the game is fair if $E=0$.

(2) Mathematical Expectation $E(X)=a_1p_1+a_2p_2+\dots\dots\dots+a_kp_k$
where the probabilities of the amounts $a_1,a_2,a_3,\dots\dots\dots,a_k$ are $p_1,p_2,p_3,\dots\dots\dots p_k$ respectively.

Example:

A player tosses 3 coins , He wins Rs.500 if 3 heads occur, Rs.300 if 2 heads occur, Rs.100 if one head occurs. On the other hand, he loses Rs.1500 if 3 tails occur. Find the value of the game to the player. Is it favourable?

Solution:

Let X= no. of head occurring in tosses of a coin.
X is a discrete random variable.

Here Sample space,

S={HHH,HHT,HTH,THH,TTH,THT,HTT,TTT}

Probability of 3 heads occurs = $P(X=3) = 1/8$.

Probability of 2 heads occurs = $P(X=2) = 3/8$.

Probability of 1 heads occurs = $P(X=1) = 3/8$.

Probability of 0 heads occurs = $P(X=0) = 1/8$.

Therefore the discrete probability distribution is

x	0	1	2	3
P(x)	1/8	3/8	3/8	1/8

Expected value of the game

$$= (500)\left(\frac{1}{8}\right) + (300)\left(\frac{3}{8}\right) + (100)\left(\frac{3}{8}\right) - (1500)\left(\frac{1}{8}\right)$$

$$= \frac{200}{8} = 25 rupess > 0$$

There fore game is favourable to the player.

Continuous Random Variable:

A random variable X which can take all possible values in a given interval is called **Continuous Random Variable**

Examples: The life of electric bulb, the volume of water flowing over a waterfall. Age, Hight, weight of individuals. Area ,Temperature, Pressure.

Probability density function:

Probability density function of a continuous random variable 'X' is a function f(x) satisfying the following conditions.

$$1.\ f(x) \geq 0$$

$$2.\ \int_{-\infty}^{\infty} f(x)dx = 1$$

$$3.\ P(a < x < b) = \int_{a}^{b} f(x)dx =$$ Area under f(x) between ordinates x=a and x=b.

Note:

(1) $P(a < x < b) = P(a \le x \le b) = P(a < x \le b) = P(a \le x < b)$

(2) Probability at a point x=a is

$$P(x = a) = \int_{a-\delta x}^{a+\delta x} f(x)dx$$

Cumulative Distribution Function:

For a continuous random variable 'X' with probability density function f(x), the cumulative distribution F(x) is defined as

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \qquad where -\infty < x < \infty$$

It follows that:

$$\Rightarrow F(-\infty) = 0, F(\infty) = 1;$$

$$\Rightarrow 0 \le F(x) \le 1$$

$$where -\infty < x < \infty$$

$$\Rightarrow f(x) = \frac{d}{dx} F(x) = F^1(x) \ge 0$$

and

$$\Rightarrow P(a < x < b) = F(b) - F(a)$$

Mean and Variance of Continuous random variable:

Mean(μ):

⇨ Mean of a distribution is given by $\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$.

⇨ If X is defined from a to b, then $\boxed{\mu = E(X) = \int_{a}^{b} x\, f(x)dx}$.

Variance (σ^2):

⇨ Variance of a distribution is given by

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \qquad \text{or} \qquad \sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2 .$$

⇨ Suppose that the variate X is defined from a to b.

$$\sigma^2 = \int_{a}^{b} (x-\mu)^2 f(x)\,dx \qquad \text{or} \qquad \sigma^2 = \int_{a}^{b} x^2 f(x)\,dx - \mu^2$$

Problem:

Suppose a continuous random variable X has the Probability density function

$$f(x) = \begin{cases} k(1-x^2) & \text{for } 0<x<1 \\ 0 & \text{otherwise} \end{cases}$$

Find (i) 'k'

(ii) $P(0.1<x<0.2)$

(iii) $P(X>0.5)$.

(iv) Using distribution function, determine the probabilities that (a) 'x' less than 0.3 i.e. $P(X<0.3)$ (b) Between 0.4 and 0.6 i.e. $P(0.4<x<0.6)$.

(v) Calculate mean and variance for the probability density function.

Solution:(i)

We know that $\int_{-\infty}^{\infty} f(x)\, dx$=1

$$\Rightarrow \int_{-\infty}^{0} f(x)\, dx + \int_{0}^{1} f(x)\, dx + \int_{1}^{\infty} f(x)\, dx = 1$$

$$\Rightarrow 0 + \int_0^1 f(x)\ dx\ + 0\, dx = 1$$

$$\Rightarrow \int_0^1 k(1-x^2)dx = 1$$

$$\Rightarrow k\left(x - \frac{x^3}{3}\right)_0^1 = 1 \Rightarrow \frac{2k}{3} = 1 \Rightarrow k = \frac{3}{2}$$

(ii)

$$P\left(0.1 < x < 0.2\right) = \int_{0.1}^{0.2} k(1-x^2)dx = \frac{3}{2}\left(x - \frac{x^3}{3}\right)_{0.1}^{0.2} = 0.1465$$

(iii)

$$\text{P(X>0.5)}=\int_{0.5}^{\infty} f(x)dx = \int_{0.5}^{1} f(x)dx + \int_{1}^{\infty} f(x)dx$$

$$= \int_{0.5}^{1} k(1-x^2)dx = \frac{3}{2}\left(x - \frac{x^3}{3}\right)_{0.5}^{1} = 0.3125 \cdot$$

(iv).We know that Cumulative distribution function

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

So

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt$$

$$= 0 + \int_{0}^{x} f(t)dt = \frac{3}{2}\int_{0}^{x}(1-t^2)dt = \frac{3}{2}\left[t - \frac{t^3}{3}\right]_0^x = \frac{3}{2}\left[x - \frac{x^3}{3}\right].$$

$$\Rightarrow P(X < 3) = F(0.3) = \frac{3}{2}\left[0.3 - \frac{0.3^3}{3}\right]$$

(iv) (a)

$$P(X < 0.3) = F(0.3) = \frac{3}{2}\left[0.3 - \frac{0.3^3}{0.3}\right].$$

(iv)(b) $P(0.4 < X < 0.6) = F(0.6) - F(0.4)$

$$= \frac{3}{2}\left[0.6 - \frac{0.6^3}{3}\right] - \frac{3}{2}\left[0.4 - \frac{0.4^3}{3}\right] = 0.224$$

(v) Mean of the function,

$$Mean = \mu$$

$$= \int_{-\infty}^{\infty} x.f(x)dx + \int_{-\infty}^{0} x.f(x)dx + \int_{0}^{1} x.f(x)dx + \int_{1}^{\infty} x.f(x)dx$$

$$= \int_0^1 x\left(\frac{3}{2}\right)[1-x^2]\text{dx} = \left(\frac{3}{2}\right)\int_0^1 x[1-x^2]dx =$$

$$\left(\frac{3}{2}\right)\int_0^1 [x-x^3]dx = \frac{3}{8}.$$

Variance of the function:

$$Variance = \sigma^2 = \int_{-\infty}^{\infty} x^2.\text{f(x)dx} - \mu^2 =$$

$$\int_0^1 x^2.\text{f(x)dx} - \mu^2 = \frac{19}{320}.$$

(Or)

$$Variance = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx == \int_0^1 \left(x-\frac{3}{8}\right)^2 \frac{3}{2}(1-x^2)dx = \frac{19}{320}$$

Example:The trouble shooting capacity of an IC chip in a circuit is a random variable X whose cumulative distribution function is given by

$$F(x) = \begin{cases} 0, \ for\ x \leq 3 \\ 1 - \frac{9}{x^2}, \ for\ x > 3 \end{cases}$$

where x denotes the number of years. Find the probability that the IC chip will work properly

(i) Less than 8 years (ii) Beyond 8 years
(iii) Between 5 to 7 years
(iv) anywhere from 2 to 5 years

Solution :

$$\text{We have } F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

$$\therefore \quad F(X) = \int_{0}^{x} f(t)\,dt \ (\because x > 0) = \begin{cases} 0, \text{if } x \leq 3 \\ 1 - \dfrac{9}{x^2}, \text{if } x > 3 \end{cases}$$

$$(i) \quad P(X \leq 8) = \int_{0}^{8} f(t)\,dt = 1 - \frac{9}{8^2} = 0.8594$$

$$(ii) \quad P(X > 8) = 1 - P(X \leq 8) = 1 - 0.8594 = 0.1406$$

$$(iii) \quad P(5 \leq x \leq 7) = F(7) - F(5) = \left(1 - \frac{9}{7^2}\right) - \left(1 - \frac{9}{5^2}\right)$$

$$= 9\left(\frac{1}{25} - \frac{1}{49}\right) = \frac{24 \times 9}{25 \times 49} = 0.1763$$

$$(iv) \quad P(2 \leq x \leq 5) = F(5) - F(2) = \left(1 - \frac{9}{5^2}\right) - 0 = \frac{16}{25} = 0.64$$

Example: Is the function defined by

$$f(x) = \begin{cases} 0, x < 2 \\ \frac{1}{18}(2x+3), 2 \le x \le 4 \\ 0, \ x > 4 \end{cases}$$

a probability density function? Find the probability that a variate having $f(x)$ as density function will fall in the interval $2 \le x \le 3$.

Solution : (*i*) For all points x in $-\infty \leq x \leq \infty, f(x) \geq 0$ and

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{2} 0.\,dx + \int_{2}^{4} \frac{1}{18}(2x+3)\,dx + \int_{4}^{\infty} 0.\,dx$$

$$= \frac{1}{18}\int_{2}^{4}(2x+3)\,dx = \frac{1}{18}\left[\frac{(2x+3)^2}{4}\right]_2^4$$

$$= \frac{1}{72}(121-49) = 1$$

Hence $f(x)$ is a probability density function.

(*ii*) The probability that the density will fall in the interval $2 \le x \le 3$ is

$$P(2 \le x \le 3) = \int_2^3 f(x)\,dx = \frac{1}{18}\int_2^3 (2x+3)\,dx$$

$$= \frac{1}{18}\left(x^2+3x\right)_2^3 = \frac{1}{18}(18-10) = \frac{8}{18} = \frac{4}{9}$$

Properties of Expectation:

1. $E(kX) = kE(X)$ **or** $\mu_{kX} = k\mu_X$

2. $E(X+k) = E(X)+k$ **or** $\mu_{X+k} = \mu_X + k$

3. $E(X+Y) = E(X)+E(Y)$ **or** $\mu_{X+Y} = \mu_X + \mu_Y$

4. Let X,Y are two random variables such that Y$\leq$ X then E(Y)$\leq$ E(X) provided expectations are exist.

- The above rules are true for both Discrete and Continuous Random variables.

Properties of variance:

$$1. Var(X+k) = Var(X)$$

$$2. Var(kX) = k^2.Var(X)$$

$$3. Var(X+Y) = Var(X) + Var(Y)$$

$$4. Var(X-Y) = Var(X) - Var(Y)$$

$$5. Var(aX+b) = a^2.Var(X)$$

$$6. Var(X) = E(X^2) - \{E(X)\}^2$$

* Similar result follows for continuous random variable X.

Lecture notes on

Probability Distributions

<u>Probability distributions are two types:</u>

1. Discrete probability distribution

2. Continuous probability distribution

- If the defined random variable is discrete then the probability distribution is called *Discrete probability Distribution* other wise (i.e. defined random variable is continuous) the probability distribution is called Continuous probability distribution.

Discrete probability distributions:

1. Binomial distribution

2. Poisson distribution.

Continuous probability distributions:

1. Normal distribution or Gaussian distribution

2. t-distribution

3. χ^2 distribution

4. F- distribution

Binomial Distribution:

Binomial distribution is a discrete probability distribution developed by James.Bernouli in 1700.

It is also known as Bernoulli distribution.

Properties of Binomial or Bernoulli distribution:

1.The outcome of each trail is classified into two mutually exclusive categories arbitrarily called a “success” and “failure”.

E.g.: tossing a coin, occurrence of head, tail are mutually exclusive.

2. Each trial is independent of other trails.

(i.e. the outcome of any observation is independent of the outcome of any other observation or trail)

3. The probability of success ‘p’ remains constant from trial to trial.

similarly the probability of failure ‘q’ or (1-p) remains constant over all observations.

4. The experiment is repeated n number of times, called n trials, where n is fixed integer.

Note:We have $(y + x)^n = nC_0\, x^n + nC_1\, x^{n-1} . y + nC_2\, x^{n-2} . y^2 + \ldots + nC_n\, y^n$

General model of Binomial distribution:

Let ‘n’ be the finite number of trials, all the trials are independent with the probility of success ‘p’ for each trial and the probability of failure is ‘q’ then

The probability of obtaining exactly ‘r’ times success out of ‘n’ trial is given by

$$f(r) = P(X = r) = nc_r . p^r . q^{n-r}$$

for r= 0,1,2,3,……..,n , and p+q=1.

- The probability of ‘x’ times success in ‘n’ trials is

$$f(x) = P(X = x) = nc_x . p^x . q^{n-x}$$

$$or$$

$$f(x) = P(X = x) = nc_x . p^x . (1 - p)^{n-x}$$

Note: The probability of ‘x’ times success in ‘n’ trials is It is denoted by B(x; n, p).

Note:

$$\Rightarrow \sum_{x=0}^{n} B(x; n, p) = 1$$

$$\Rightarrow nc_0 . p^0 . q^n + nc_1 . p^1 . q^{n-1} + nc_2 . p^2 . q^{n-2} + \dots\dots\dots + nc_n . p^n . q^0 = 1$$

Mean of Binomial distribution:

$$mean = \mu = E(x) = \sum_{x=0}^{n} x.p(x)$$

$$\mu = E(x) = \sum_{x=0}^{n} x.B(x;n,p)$$

$$= \sum_{x=1}^{n} x.\frac{n!}{(n-x)!x!}.p^{x}.q^{n-x}$$

$$= np\sum_{x=1}^{n} \frac{(n-1)!}{(n-x)!(x-1)!}.p^{x-1}.q^{n-x}$$

$$= np\left[\frac{(n-1)!}{(n-1)!0!}.p^{0}.q^{n-1} + \frac{(n-1)!}{(n-2)!1!}.p^{1}.q^{n-2} + \frac{(n-1)!}{(n-3)!2!}.p^{2}.q^{n-3} + \dots\dots\dots\dots + \frac{(n-1)!}{(n-n)!(n-1)!}.p^{n-1}.q^{0}\right]$$

$$= np\left[(p+q)^{n-1}\right]$$

$$= np$$

Variance of Binomial distribution:

$$\sigma^2 = E(x^2) - \{E(x)\}^2 = \sum_{x=0}^{n} x^2.p(x) - \left[\sum_{x=0}^{n} x.p(x)\right]^2 = \sum_{x=0}^{n} x^2.p(x) - \mu^2$$

$$consider, E(x^2) = \sum_{x=0}^{n} x^2.B(x;n,p)$$

$$= \sum_{x=1}^{n} (x^2 - x + x).\frac{n!}{(n-x)!x!}.p^x.q^{n-x}$$

$$= \sum_{x=2}^{n} x(x-1)\frac{(n-1)!}{(n-x)!(x-1)!}.p^{x-1}.q^{n-x} + \sum_{x=1}^{n} x.\frac{n!}{(n-x)!x!}.p^x.q^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n!}{(n-x)!(x-2)!}.p^x.q^{n-x} + \mu$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(n-x)!(x-2)!}.p^{x-2}.q^{n-x} + \mu$$

$$= n(n-1)p^2 \begin{bmatrix} \frac{(n-2)!}{(n-2)!0!}.p^0.q^{n-2} + \frac{(n-2)!}{(n-3)!1!}.p^1.q^{n-3} + \frac{(n-2)!}{(n-4)!2!}.p^2.q^{n-4} + ... \\ ....+ \frac{(n-2)!}{(n-n)!(n-2)!}.p^{n-2}.q^0 \end{bmatrix} + \mu$$

$$= n(n-1)p^2 \left[(p+q)^{n-2} \right] + np$$

$$= n(n-1)p^2 + np$$

$$therefore, \sigma^2 = n(n-1)p^2 + np - (np)^2 = np - np^2 = np(1-p) = npq$$

$$\sigma^2 = npq$$

$$S.D. = \sqrt{\sigma^2} = \sqrt{npq}$$

Problems:(1) A fair coin is tossed 6 times, find the probability of getting

(i) exactly two heads

(ii) no head

(iii) at least 4 heads

(iv) at least one head

<u>Solution:</u> Given number of trails n=6,

p=P(success)= probability of getting head= $\frac{1}{2}$

q=P(failure)= probability of getting tail= $\frac{1}{2}$

$$(i) P(X=2) = 6C_2 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^{6-2} = 6C_2 \cdot \left(\frac{1}{2}\right)^6 = \frac{15}{64}.$$

(ii): $\mathrm{P(No\ Head)} = \mathrm{p}(X = 0) \quad = 6C_0 \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^6 = \frac{1}{64}$

$$(iii) P(X \geq 4) = P(X=4) + P(X=5) + P(X=6)$$

$$= 6C_4 \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^2 + 6C_5 \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^1 + 6C_6 \cdot \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^0$$

$$= \left(6C_4 + 6C_5 + 6C_6\right) \cdot \left(\frac{1}{2}\right)^6$$

$$(iv) P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6)$$

(OR) $\quad P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 63/64$

(2) A fair die is thrown 7 times. Determine the probability that 5 or 6 appears
(i) exactly 3 times
(ii) never occurs

Solution: Given n=7,
probability of success=p=P(Getting 5or 6)=1/6+1/6=2/6=1/3.

probability of failure=q=2/3.

$$(i)P(X=3)=7C_3.\left(\frac{1}{3}\right)^3.\left(\frac{2}{3}\right)^{7-3}=7C_3.\left(\frac{16}{3^7}\right)=\frac{560}{2187}=0.2560$$

$$(ii)P(X=0)=7C_0.\left(\frac{1}{3}\right)^0.\left(\frac{2}{3}\right)^7=\left(\frac{2^7}{3^7}\right)=\frac{128}{2187}=0.05852$$

Problem(3): In a sampling a large number of parts manufactured by a machine, the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples, how many would be expected to contain at least 3 defective parts?

Solution:

Given n=20,mean μ=2.

Let X be the number defective items

$$mean = \mu = np$$

$$\Rightarrow 20p = 2$$

$$\Rightarrow p = \frac{2}{20} = \frac{1}{10} = 0.1,$$

$$then$$

$$q = 1 - p = 1 - 0.1 = 0.9$$

$$P(atleast3defectiveparts) = P(X \geq 3)$$

$$\Rightarrow P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - \{P(X = 0 + P(X = 1) + P(X = 2)\}$$

$$= 1 - \left[\left(20C_0.(0.1)\right)^0.(0.9)^{20} + \left(20C_1.(0.1)\right)^1.(0.9)^{19} + \left(20C_2.(0.1)\right)^2.(0.9)^{18}\right]$$

$$= 0.323$$

The expected number of samples containing at least 3 defective items = 1000(0.323)

=323 samples.

Recurrence relation or Recursive formula for B.D.

$$B(r:n,p) = nc_r.p^r.q^{n-r}$$

$$P(X=r) = \frac{n!}{(n-r)!r!}.p^r.q^{n-r} \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_(1)$$

$$B(r+1:n,p) = nc_{r+1}.p^{r+1}.q^{n-(r+1)}$$

$$P(X=r+1) = \frac{n!}{(n-(r+1))!(r+1)!}.p^{r+1}.q^{n-(r+1)} \_\_\_\_\_\_\_\_\_\_(2)$$

$$\frac{(2)}{(1)} \Rightarrow \frac{P(r+1)}{P(r)} = \frac{\frac{n!}{(n-(r+1))!(r+1)!}p^{r+1}.q^{n-(r+1)}}{\frac{n!}{(n-r)!r!}p^r.q^{n-r}} = \frac{(n-r)}{r+1}.\frac{p}{q}$$

$$\Rightarrow P(r+1) = \frac{(n-r)}{r+1}.\frac{p}{q}.P(r)$$

Problem: The ratio of the probabilities of 3 success and 2 success among 5 independent trials is 1/3. Find the probility of getting success 'p'?

Solution:

$$P(X = r) = nc_r . p^r . q^{n-r}$$

$$\Rightarrow \frac{P(X = 3)}{P(X = 2)} = \frac{1}{3}$$

$$\Rightarrow \frac{5c_3 . p^3 . q^2}{5c_2 . p^2 . q^3} = \frac{1}{3}, \{5c_3 = 5c_2\}$$

$$\Rightarrow \frac{p}{q} = \frac{1}{3}$$

$$\Rightarrow \frac{p}{1-p} = \frac{1}{3}$$

$$\Rightarrow p = \frac{1}{4}$$

Example:

Find the maximum n such that the probability of getting no head in

tossing a fair coin n times is greater than 0.1.

Solution : p = The probability of getting a head = $\frac{1}{2}$

q = The probability of not getting a head = $1 - p = \frac{1}{2}$

Given that $P(X = 0) > 0.1$

$i.e., \ {}^nC_0 \ p^0 q^n > 0.1 \qquad i.e., \ q^n > 0.1$

$i.e., \left(\frac{1}{2}\right)^n > 0.1 \Rightarrow 2^n < 10 \Rightarrow n < 4.$ So $n = 3$

Example:

20% of items produced from a factory are defective. Find the probability

that in a sample of 5 chosen at random

(i) none is defective (ii) one is defective (iii) $p\ (1 < x < 4)$

Solution : Probability of defective items = $p = 20\% = 0.2$

Probability of non defective items = $q = 1 - p = 1 - 0.2 = 0.8$

Total number of items, $n = 5$

(i) Probability that none is defective = Probability of 0 defective item

$$= p(0) = {}^5C_0\ (0.2)^0\ (0.8)^5 = (0.8)^5 = 0.32768$$

(ii) Probability of 1 defective item $= p(1) = {}^5C_1\ (0.2)^1\ (0.8)^4$

$$= 5(0.2)\ (0.4096) = 0.4096$$

$(iii)\quad p\,(1 < x < 4) = p(2) + p(3)$

$$= {}^5C_2\,(0.2)^2\,(0.8)^3 + {}^5C_3\,(0.2)^3\,(0.8)^2$$

$$= (0.2)^2\,(0.8)^2\,[10(0.8) + 10\,(0.2)]$$

$$= 0.0256\,(8 + 2) = 0.256$$

<u>Poisson Distribution:</u>

If the parameters ‘n’ and ‘p’ of a Binomial distribution are known, where ‘n’ is very large and ‘p’ is very small then the application of *<u>Binomial distribution</u>* becomes difficult.

How ever if we assume that as $n \to \infty, and \ \ p \to 0$ such that ‘np’ always remains finite (say λ), we get ***<u>Poisson Approximation of binomial distribution.</u>***

Let np=λ , then p=λ/n

From binomial distribution we have

$$P(X=r)=\frac{n!}{(n-r)!r!}.p^{r}.q^{n-r}$$

$$=\frac{n!}{(n-r)!r!}.p^{r}.(1-p)^{n-r}$$

$$=\frac{n(n-1)(n-2)............(n-(r-1))(n-r)!}{(n-r)!r!}.\left(\frac{\lambda}{n}\right)^{r}.\left(1-\frac{\lambda}{n}\right)^{n-r}$$

$$= \frac{n(n-1)(n-2)............(n-(r-1))}{r!} \cdot \left(\frac{\lambda}{n}\right)^r \cdot \left(1-\frac{\lambda}{n}\right)^{n-r}$$

$$= \left(\frac{\lambda^r}{r!}\right) \cdot \frac{n(n-1)(n-2)............(n-(r-1))}{n^r} \cdot \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$as, n \to \infty$

$$= \left(\frac{\lambda^r}{r!}\right) \cdot \left(1-\frac{1}{n}\right) \cdot \left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)................\left(1-\frac{r-1}{n}\right) \cdot \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$\left(\begin{array}{l} as, n \to \infty \\ \left(1-\frac{1}{n}\right) \cdot \left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)................\left(1-\frac{r-1}{n}\right) \to 1 \\ and \\ \left(1-\frac{\lambda}{n}\right)^r \to 1 \end{array}\right)$$

$$\therefore P(X=r)=\frac{\lambda^r}{r!}\cdot\left(\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\right),$$

$$\left[we\,have, \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x\right]$$

$$\Rightarrow P(X=r)=\frac{\lambda^r}{r!}.e^{-\lambda},\ for$$ r =0,1,2,3…………..

This probability distribution is called Poisson Probability Distribution.

Poisson Probability Distribution:

A discrete Random variable X is said to follow Poisson Probability Distribution with parameter λ if its probability mass function is

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

(OR)

A discrete Random variable X is said to follow Poisson Probability Distribution with parameter λ if its probability mass function is

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

Note: λ is the parameter of the distribution.

Note:The sum of the Poisson probabilities for r=0,1,2,…..is ‘1’.

i.e.

$$p(0)+p(1)+p(2)+.........=e^{-\lambda}+e^{-\lambda}.\frac{\lambda}{1!}+e^{-\lambda}.\frac{\lambda^2}{2!}+e^{-\lambda}.\frac{\lambda^3}{3!}+.........$$

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

$$=e^{-\lambda}.e^{\lambda}=e^{0}=1$$

Mean and variance of Poisson distribution:

$$mean = \mu = E(X) = \sum_{r=0}^{\infty} r.P(r)$$

$$= \sum_{r=0}^{\infty} r.\left(\frac{e^{-\lambda}.\lambda^{r}}{r!}\right)$$

$$= e^{-\lambda} \sum_{r=1}^{\infty} \left(\frac{\lambda^{r}}{(r-1)!}\right)$$

$$= e^{-\lambda} \left[\frac{\lambda}{0!} + \frac{\lambda^{2}}{1!} + \frac{\lambda^{3}}{2!} + \frac{\lambda^{4}}{3!} + \dots\dots\dots\dots\right]$$

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

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

$$= \lambda$$

$$\therefore mean = \mu = E(X) = \lambda$$

$$\sigma^2 = E(X^2) - \left\{E(X)\right\}^2$$

$$= \sum_{r=0}^{\infty} r^2 . \left(\frac{e^{-\lambda} . \lambda^r}{r!} \right) - \lambda^2$$

$$= e^{-\lambda} \sum_{r=1}^{\infty} \frac{r^2 . \lambda^r}{r!} - \lambda^2$$

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

$$= e^{-\lambda} . \lambda \left[\frac{1}{1!} + 2\lambda + 3 . \frac{\lambda^2}{2!} + 4 . \frac{\lambda^3}{3!} + \ldots\ldots\ldots \right] - \lambda^2$$

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

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

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

$$= e^{-\lambda}.\lambda\left\{e^{\lambda} + \lambda e^{\lambda}\right\} - \lambda^2$$

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

$$= \lambda + \lambda^2 - \lambda^2$$

$$= \lambda$$

$$\therefore \sigma^2 = \lambda$$

Recurrence relation for Poisson Distribution:

We have

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

$$P(X = r+1) = \frac{e^{-\lambda}.\lambda^{r+1}}{(r+1)!}$$

$$\frac{P(r+1)}{P(r)} = \frac{\dfrac{e^{-\lambda}.\lambda^{r+1}}{(r+1)!}}{\dfrac{e^{-\lambda}.\lambda^{r}}{r!}} = \frac{\lambda}{r+1}$$

$$P(r+1) = \frac{\lambda}{r+1}.P(r)$$

Problems:(1) Determine the probability that 2 of 100 books bound will be defective if it is known that 5% of books bound at this bindery are defective.

(i) Use Binomial distribution

(ii) Use Poisson approximation to binomial distribution

Solution:

$$n = 100$$

$$p = 5\% defective = 0.05$$

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

$$(i) P(X = 2) = nC_2.(p)^2.(q)^{n-2} = 100C_2.(0.05)^2.(0.95)^{100-2} = 0.081$$

$$(ii) \lambda = np = 100 \times (0.05) = 5$$

$$\Rightarrow P(X = 2) = \frac{\lambda^2.e^{-\lambda}}{2!} = \frac{5^2.e^{-5}}{2!} = \frac{25.e^{-5}}{2} = 0.084$$

Example: Wireless sets are manufactured with 25 soldered joints each. Probability of a joint is defective 1/500 . How many sets can be expected to be free from defective joints in a consignment of 10000 sets?

Solution :

Number of soldered joints = 25 = n

Probability that a joint is defective = $\frac{1}{500}$

$$\therefore \quad \text{Mean} = np = 25 \times \frac{1}{500} = \frac{1}{20} = .05$$

Thus $\lambda = .05$

$$\text{P(There are } r \text{ defective joints)} = \frac{e^{-\lambda}.\lambda^{r}}{r!}$$

$$\therefore \quad P(r = 0) = e^{-\lambda} = e^{-.05}$$

Hence expected number of sets free of defective joints among 10000 sets

$$= 10000 \,.\, e^{-.05} \quad = 10000\,(0.951229) = 9512.29 \approx 9512$$

Example: If 2% of light bulbs are defective. Find

(i) At least one is defective.

(ii) Exactly 7 are defective

(iii) $p(1 < x < 8)$ in a sample of 100

<u>Solution :</u> Given $n = 100$ and

$p =$ The probability of defective light bulbs = 0.02

$\therefore$ Mean $= \lambda =$ Mean number of defective light bulbs in a sample of 100.

$$= np = 100\,(0.02) = 2$$

Since p is small, we use Poisson distribution.

Probability of x defective light bulbs in a sample of 100 is

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

(*i*) p (at least one is defective) $= p(x \geq 1) = 1 - p(x = 0)$

$$= 1 - \frac{e^{-2} . 2^0}{0!} = 1 - e^{-2} = 0.8646$$

(*ii*) p (exactly 7 are defective) $= p(x = 7)$

$$= \frac{e^{-2} . 2^7}{7!} = \frac{128\, e^{-2}}{5040} = 0.0034$$

(*iii*) $p(1 < x < 8) = p(x = 2) + p(x = 3) + p(x = 4) + p(x = 5) + p(x = 6) + p(x = 7)$

$$= e^{-2}\left[\frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!} + \frac{2^5}{5!} + \frac{2^6}{6!} + \frac{2^7}{7!}\right]$$

$$= e^{-2}\left[\frac{4}{2} + \frac{8}{6} + \frac{16}{24} + \frac{32}{120} + \frac{64}{720} + \frac{128}{5040}\right]$$

$$= e^{-2}(2 + 1.33 + 0.66 + 0.27 + 0.09 + 0.02)$$

$$= e^{-2}(4.37) = 0.59$$

Example: **Given that P(X=2)=9P(X=4)+90 P(X=6) for Poisson variate X. Find (i)P(X<2), (ii)P(X>4), (iii)P(X≥1).**

Solution : If X is a Poisson variate with parameter λ , then

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

Since $P(X=2)=9P(X=4)+90P(X=6)$, we have

$$\frac{e^{-\lambda}.\lambda^{2}}{2!}=9.\frac{e^{-\lambda}.\lambda^{4}}{4!}+90.\frac{e^{-\lambda}.\lambda^{6}}{6!}$$

$$=\frac{e^{-\lambda}.\lambda^{2}}{8}[3\lambda^{2}+\lambda^{4}]$$

or $4 = 3\lambda^2 + \lambda^4$ or $\lambda^4 + 3\lambda^2 - 4 = 0$ or $(\lambda^2 + 4)(\lambda^2 - 1) = 0$

$\therefore\ \lambda^2 + 4 = 0$ or $\lambda^2 - 1 = 0$. But λ cannot be imaginary.

[(or) This is a quadratic equation in λ^2

$$\therefore\ \lambda^2 = \frac{-3 \pm \sqrt{9+16}}{2} = \frac{-3 \pm 5}{2} = 1, -4 \Bigg]$$

Since $\lambda > 0$, we get

$\lambda^2 = 1 \Rightarrow \lambda = 1,$ taking positive sign

Hence $P(X = x) = \dfrac{e^{-1}.1^x}{x!}, x = 0, 1, 2, ...$

(i) $P(x < 2) = P(x = 0) + P(x = 1)$

$$= e^{-1} + e^{-1} = \frac{2}{e} = 0.7358$$

(ii) $P(x > 4) = 1 - P(x \le 4)$

$$= 1 - [P(x = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4)]$$

$$= 1 - \left[e^{-1} + e^{-1} + \frac{e^{-1}}{2} + \frac{e^{-1}}{3!} + \frac{e^{-1}}{4!}\right]$$

$$= 1 - e^{-1}\left[2 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24}\right] = 1 - e^{-1}\left[\frac{65}{24}\right]$$

$$= 1 - 0.9963 = 0.0037$$

(iii) $P(x \ge 1) = 1 - p(x = 0) = 1 - e^{-1} = 1 - 0.3679 = 0.6321$

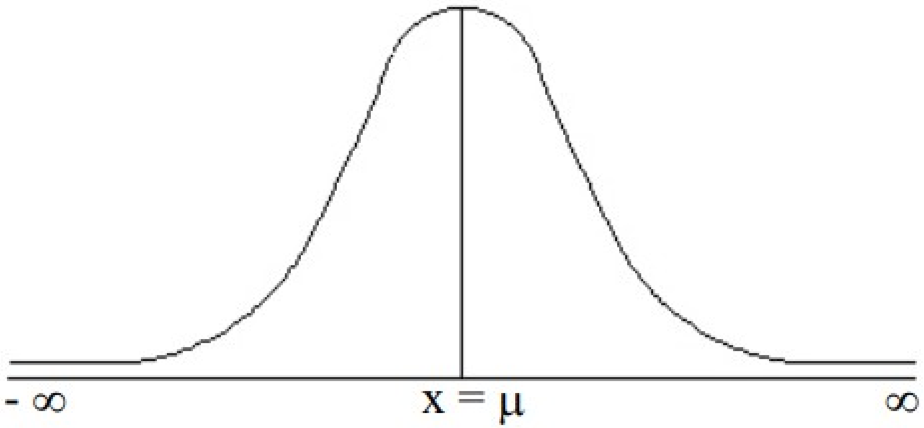
Normal Distribution or Gaussian Distribution:

A random variable X is said to have Normal probability distribution or simply normal distribution if its probability Density function is given by

.

Here μ= mean, σ=standard deviation, are two parameters of continuous random variable X.

Normal probability curve: The curve representing the normal distribution is called the normal probability curve.



Mean

The average value

How to find the Mean:

1. Add up all the numbers.
2. Divide the sum by the number of values.

E.g. The mean of 3,2,10,5 is

$$\frac{3+2+10+5}{4}=\frac{20}{4}=5$$

Median

The middle number

How to find the Median:

1. Put the numbers from smallest to largest.
2. The number in the middle is the median. If there are two middle numbers, add them and divide by two.

Mode

The most frequent number

Special Cases:

- No Mode if all the numbers occur the same amount of times.
- More than one Mode if more than one number is the most frequent.

Range

Difference between highest and lowest numbers

How to find the Range:

1. Put the numbers from smallest to largest.
2. Subtract the lowest value from the largest.

Characteristics of Normal distribution Curve:

1. The graph of the Normal distribution y= f(x) in XY-plane is bell shaped and symmetrical about the line x=μ.

2. The maximum value of Gaussian density function is

$\frac{1}{\sqrt{2\pi\sigma^2}}$(it will occur at x= μ)

3. The mean, median and mode are coincide and therefore Normal curve has only one maximum point(i.e N.C is unimodal).

4. Normal curve has inflection points at $\mu \pm \sigma$

5. Normal curve is asymptotic to both positive X-axis and negative X-axis.

6. Area under the Normal curve is unity.

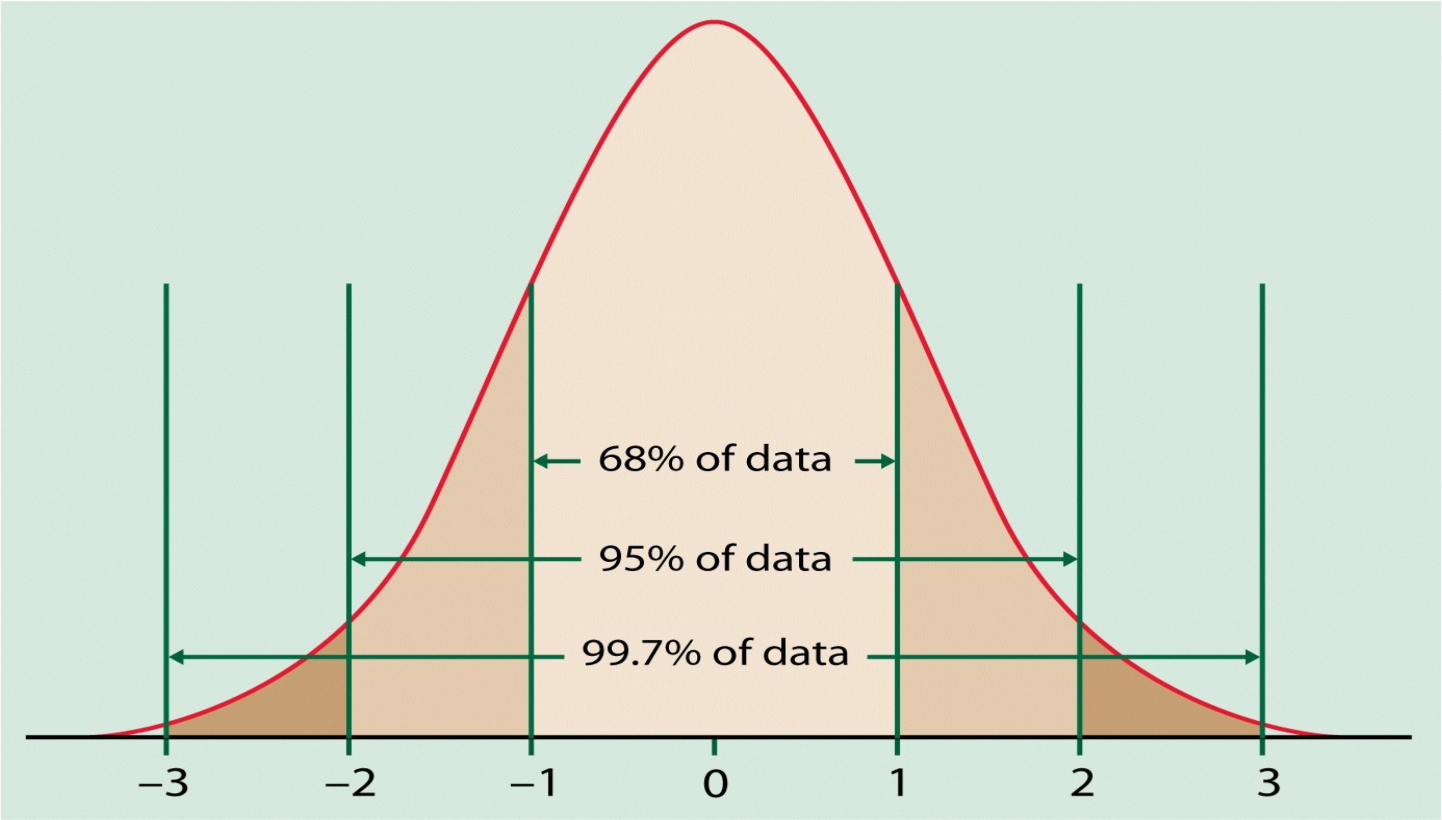
7. Change of scale from X-axis to Z-axis.

$$P\left(x_1 \le X \le x_2\right) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} . e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

8. N(0,1) is known as *Standard Normal Distribution* with mean ‘0’ and standard deviation ‘1’.

9. Area under the Normal curve is distributed as follows

–



■The standard normal distribution has mean = 0 and standard deviation sigma=1.

Area Property

= 0.6826

= 0.9544

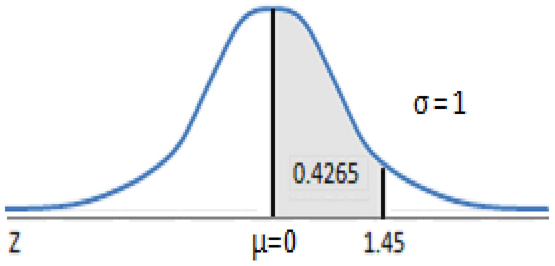
= 0.9973

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 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 .

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

Area under the standard normal curve from 0 to z.

This table provides the area between the mean and some Z score.
For example, when Z score = 1.45 the area = 0.4265.



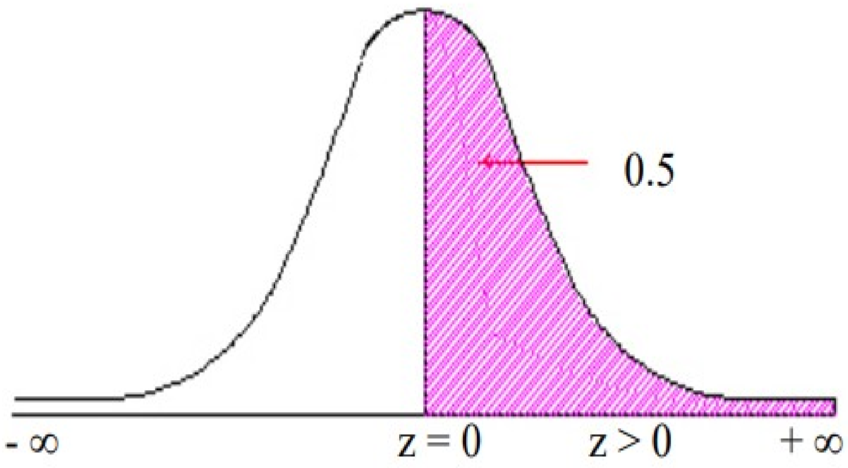
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767

2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Exmple: Students of a class were given an aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percentage of students scored. i) More than 60 marks (ii) Less than 56 marks (iii) Between 45 and 65 marks.

Solution: Given 5.

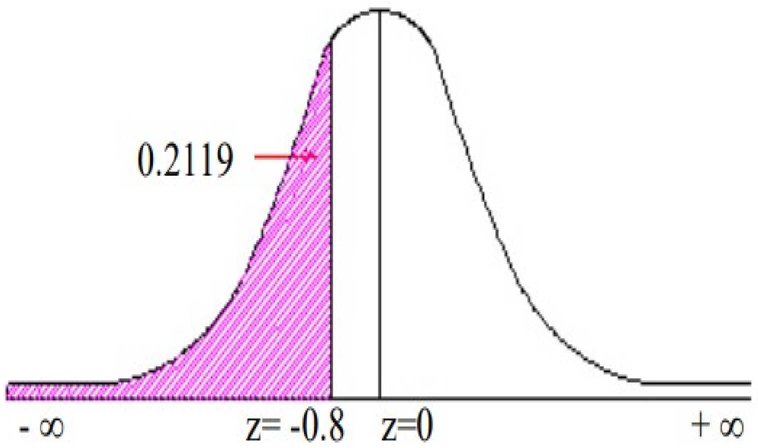
The standard normal variate, Z



(i) If x = 60,
P

Hence the percentage of students scored more than 60 marks is 0.5(100) = 50 %

(ii) If x = 56,
P

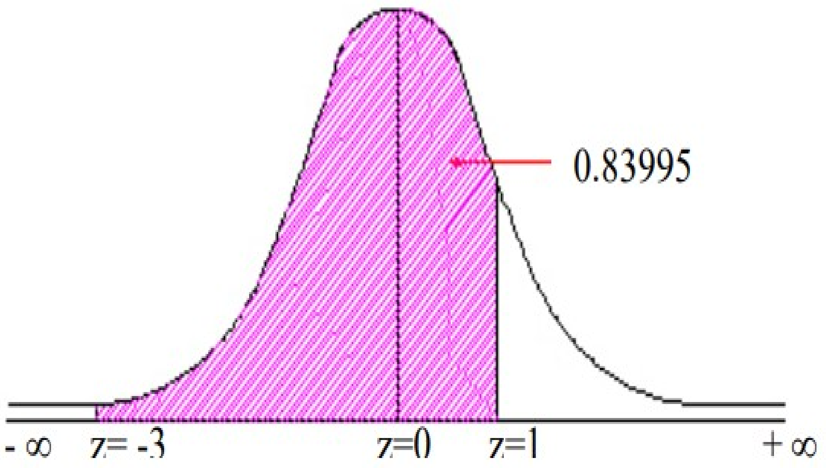


Hence the percentage of students score less than 56 marks is 0.2119(100) = 21.19 %

(iii) If , .

, 1

P



= 0.4986 + 0.3413 (from the table) = 0.8399 Hence the percentage of students scored between 45 and 65 marks is 0.8399(100) = 83.99 %

Example: For normally distributed variate with mean 1 and standard deviation 3,Find the probabilities that (i)$3.43 \leq x \leq 6.19$ (ii) $-1.43 \leq x \leq 6.19$.

Solution : Given $\mu = 1$ and $\sigma = 3$

(*i*) When $x = 3.43$,

$$z = \frac{x-\mu}{\sigma} = \frac{3.43-1}{3} = \frac{2.43}{3} = 0.81 = z_1 \text{ (say)}$$

When $x = 6.19$,

$$z = \frac{x-\mu}{\sigma} = \frac{6.19-1}{3} = \frac{5.19}{3} = 1.73 = z_2 \text{ (say)}$$

$P(3.43 \le X \le 6.19)$

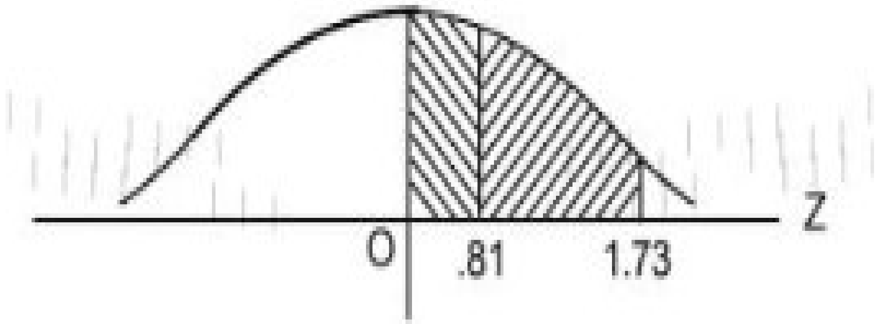
$= P(0.81 \le Z \le 1.73)$

$= P(0 \le Z \le 1.73) - P(0 \le Z \le 0.81)$

=(Area under the curve from Z=0 t0 Z=1.73)-(Area under the curve from Z=0 t0 Z=0.81)

=(1.73)-A(0.81)

=0.4582-0.2910=0.1672



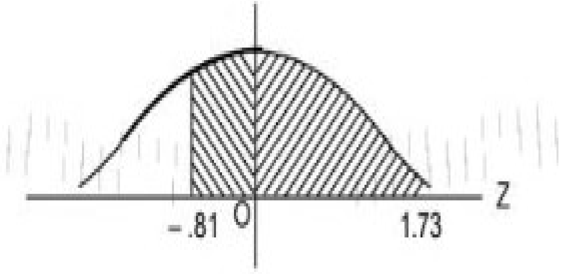
(*ii*) When $x = -1.43$,

$$z = \frac{x-\mu}{\sigma} = \frac{-1.43-1}{3} = -0.81 = z_1 \text{ (say)}$$

When $x = 6.19$,

$$z = \frac{x-\mu}{\sigma} = \frac{6.19-1}{3} = 1.73 = z_2 \text{ (say)}$$

$P(-1.43 \leq X \leq 6.19)$
$= P(-0.81 \leq Z \leq 1.73)$
$= P(-0.81 \leq Z \leq 0) + P(0 \leq Z \leq 1.73)$
$= P(0 \leq Z \leq 0.81) + P(0 \leq Z \leq 1.73)$
=(Area under the curve from Z=0 t0 Z= =0.81)+(Area under the curve from Z=0 t0 Z 1.73)
$= A(1.73) + A(0.81)$
$= 0.4582 - 0.2910 = 0.7492$



Example:

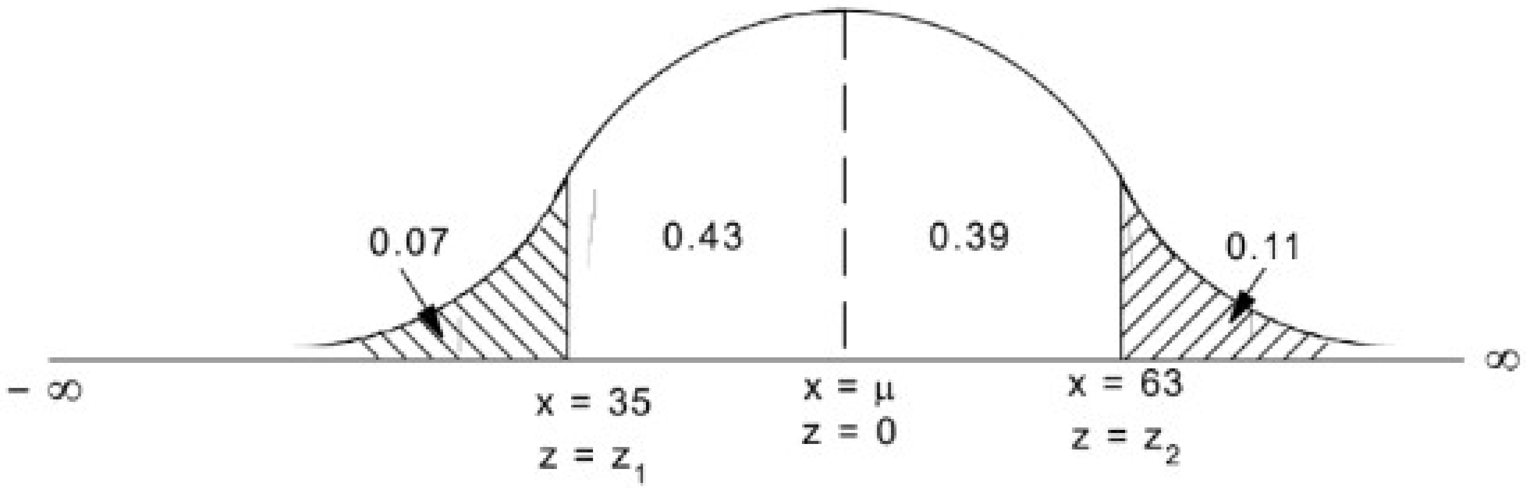
In a normal distribution 7 % of the items are under 35 and 89% of items are under 63. Find the mean and variance of the distribution.

Solution : Let μ be the mean (at $z = 0$) and σ the standard deviation of the normal curve. 7% of the items are under 35 means the area to the left of the ordinate $x = 35$.

Given $P(X < 35) = 0.07$ and $P(X < 63) = 0.89$

$\therefore \quad P(X > 63) = 1 - P(X < 63) = 1 - 0.89 = 0.11$

The points $X = 35$ and $X = 63$ are shown in the following Figure.



When x = 35, $z = \dfrac{x-\mu}{\sigma} = \dfrac{35-\mu}{\sigma} = -z_1$ (say) ... (1)

When x = 63, $z = \dfrac{x-\mu}{\sigma} = \dfrac{63-\mu}{\sigma} = z_2$ (say) ... (2)

From the figure, we have

$P(0 < z < z_2) = 0.39 \Rightarrow z_2 = 1.23$ (from Normal tables)

and $P(0 < z < z_1) = 0.43 \Rightarrow z_1 = 1.48$

From (1), we have $\dfrac{35-\mu}{\sigma} = -1.48$ .. (3)

From (2), we have $\dfrac{63-\mu}{\sigma} = 1.23$ ... (4)

(4) – (3) gives

$$\frac{28}{\sigma} = 2.71 \Rightarrow \sigma = \frac{28}{2.71} = 10.332$$

From (3), $35 - \mu = -1.48\,(\sigma) = (-1.48)\,(10.332) = -15.3$

$\therefore\ \mu = 35 + 15.3 = 50.3$

and variance $= \sigma^2 = 106.75$

Problem:

X is normal distribution with mean 2 and standard deviation 3. Find the value of the variable x such that the probability of the interval from mean to that value is 0.4115?

Solution: Given .

Suppose z_1 is required standard value, Thus $P(0 < z < z_1) = 0.4115$.

From the table the value corresponding to the area 0.4115 is 1.35 that is $z_1 = 1.35$.

Here

Thank you