Unit - I

Elements of Set Theory:

A set is a well defined collection or aggregate of objects having given properties and specified according to a well defined rule. In other words, a set is a collection of well defined objects. By well defined objects, we mean thereby that given any objects, It is possible to determine whether the object is objects, we mean thereby that given any object, it is possible to determine whether the object is a member of the set or not. The objects of a set are called elements or members. Generally, a set is denoted by capital letters - A, B, C, etc. and the elements are denoted by small letters - a, b, c, etc,

Sample Space:

The set or aggregate of all possible outcomes is knowns as sample space, For example, when we roll a die, the possible outcomes are 1, 2, 3, 4, 5 and 6 ; one and ~~of~~ only one face come upwards. Thus, all the outcomes - 1, 2, 3, 4, 5 and 6 are sample space. And each possible outcome or element in a sample space called sample point.

Probability:-

When the probability of an event is absolutely certain, the probability is said to be unity or 1. Every person is certain to die one day hence the probability is equal to unity. Mathematicians will say that the probability of man's death is one. Similarly the probability of man surviving without blood is 0. Thus, a probability can be 1 and 0. In between 1 and 0, there are fractions denoting the probabilities of all sort of events occuring. Similarly, if we toss a coin, the probability of head up or tail up is unity because it will not stand on its edge. The probability of head coming up is 1/2 and tail coming up 1/2.

Addition Theorem:-

The Simplest and most important rule used in the calculation is the addition rules, it states, "If two events are mutually exclusive, then the probability of the occurance of either A or B is the sum of the probability of A and B. Thus,

$$P(A \text{ or } B) = P(A) + P(B)$$

Example:- A bag contains 4 white, 3 black and 5 red balls, What is the probability of getting a white or red ball at random in a single draw?

Solution:- A bag contains 4 white, 3 black and 5 red balls.

The probability of getting a white ball = $4/12$

The probability of getting a red ball = $5/12$

The probability of a white or a red = $4/12 + 5/12$

$= 9/12$

$9/12 \times 100 = 75\%$

When events are not mutually exclusive:-

The addition theorem studied above is not applicable when the events are not mutually exclusive. In such cases where the events are not mutually exclusive, the probability is:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Example:- Two students x and y work independently on a problem. The probability that x will solve it is 3/4 and the probability that y will solve it is 2/3. What is the probability that the problem will be solved?

~~Sol~~ **Solution:-**

$$P(A \text{ or } B) = P(A) + P(B) - (P(A \text{ and } B)$$

The probability that x will solve the problem $= 3/4$

The probability that y will solve the problem $= 2/3$

The events are not mutually exclusive as both of them may solve the problem.

Therefore, the probability $= 3/4 + 2/3 - (3/4 \times 2/3)$

$$= \frac{17}{12} - \frac{6}{12} = \frac{11}{12}$$

Alternatively :- The probability that x will solve it and y fail to solve it $= 3/4 \times 1/3 = 3/12$

Probability that the problem will be solved $= 2/3 + 3/12$

$= 11/12$

<u>Alternatively</u> :- The probability that x will fail to solve and y ~~fail~~ solve it. $= 1/4 \times 2/3 = 2/12$

Alternatively :- The probability that neither X nor Y will solve it $1/4 \times 1/3 = 1/12$.

Hence, the probability that the problem will be solved $= 1 - 1/12 = 11/12$.

<u>Multiplication</u> :-

When it is desired to estimate the changes of the happening of successive events, the separate probabilities of these successive events are multiplied. If two events A and B are independent, then the probability that both will occur is equal to the product of the respective probabilities. We find the probability of the happening of two or more events in succession. Symbolically:-

$$P(A \text{ and } B) = P(A) \times P(B)$$

Example 1:-

In two tosses of a fair coin, what are the chances of head in both?

Solution:-

Probability of head in first toss = $1/2$

Probability of head in the second toss = $1/2$

Probability of head both tosses = $1/2 \times 1/2 = 1/4$.

Probability of head

Example 2:-

The probability that x and y will be alive ten years hence is 0.5 and 0.8 respectively. What is the probability that both of them will be alive ten years hence?

Solution:-

Probability of x being alive ten years hence = 0.5

Probability of y being alive ten years hence = 0.8

Probability x and y both being alive ten years hence = 0.5×0.8

$= 0.4.$

Simple problems:-

*. On a six-sided die, each side has a number between 1 and 6. What is the probability of throwing a 3 or a 4?

Ans:- On a six-sided die, the probability of throwing any number is 1 in 6. The probability of throwing a 3 or a 4 is double that, or 2 in 6. This can be simplified by dividing both 2 and 6 by 2.

Therefore, the probability of throwing either a 3 or 4 is 1 in 3.

*. Three coins are tossed up in the air, one at a time. What is the probability that two of them will land heads up and one will land tails up?

Ans:- HHH HHT HTT HTH TTT TTH THT THH

Notice that out of the 8 possible outcomes, only 3 of them meet that desired condition that two coins land heads up and one coin lands tails up. Probability, by definition, is the number of desired outcomes divided by the number of possible outcomes. Therefore, the probability of two heads and one tails is 3/8. ~~Choice~~.

$= 3/8.$

Unit - II

Binomial distribution, Poisson Distribution and Normal Distribution - Concept, definition, properties (without proof) and uses, simple problems in Binomial Distribution only.

Theoretical Distribution :-

In the previous chapters we have dealt with the observed frequency distribution. Broadly speaking, the frequency distributions are of two types: Observed Frequency Distribution and Theoretical Frequency Distribution. The following are important distributions :-

⇒ Binomial Distribution
⇒ Poisson Distribution
⇒ Normal Distribution

BINOMIAL DISTRIBUTIONS :-

This distribution was discovered by a Swiss mathematician, James Bernouli (1654 - 1705) and is also known as Bernoulli Distribution. He discovered this theory and published it in the year 1700 dealing with dichotomous classification of events, one processing and the other not possessing. The distribution can be used under the following conditions:-

1. The number of trials is finite and fixed.

2. In every trial, there are only two possible outcomes - success or failure.

3. The trial, there are only two are independent. The outcome of one trial does not affect the other trial.

4. p, the probability of success from trial, is fixed and q. the probability of failure, is equal to $1-p$. This is the same in all the trials.

<u>Example</u>: A head or a tail can be had on a toss of coin; a card drawn may be black or red; and item inspected from a batch may be defective or non-defective. In each experi-ment the outcome can be classified as success or failure. Success is generally denoted by p and failure is $1-p=q$.

If a single coin is tossed, the outcomes are two: head or tail. Probability of head is $1/2$ and tail is $1/2$.

Thus $(q+p)^n = (1/2 + 1/2)^1 = 1$.

If two coins are thrown the outcomes are four:

HH	TH	HT	TT
pp	qp	pq	qq
p^2	$2pq$		q^2

If three coins are tossed, the following are the outcomes:

$(p+q)^3 = p^3 + 3p^2q + 3q^2p + q^3$.

Scanned by TapScanner

PROPERTIES OF BINOMIAL DISTRIBUTION :-

1. Binomial distribution has two parameters- n and p (or q)

2. Mean = (np)

3. Varience = npq

4. Standard deviation = $\sqrt{npq}$

5. Skewness $(\beta_1) = \frac{(q-p)^2}{npq}$

6. Kurtosis $(\beta_2) = 3 + \frac{1-6pq}{npq}$

7. Binomial distribution is symmetrical if $p = q = 0.5$

8. It is positively skewed if $p < 0.5$ and it is negatively skewed if $p > 0.5$.

Example :-

Five coins are tossed 3,200 times, Find the frequencies of the distribution heads and tails and tabulate the results.

Solutions :-

$p = 0.5$ and $q = 0.5$

Applying binomial distribution, the probability of getting x heads is given by :

$$P(x) = {}^5C_x \left(\tfrac{1}{2}\right)^x \left(\tfrac{1}{2}\right)^{5-x} = {}^5C_x \left(\tfrac{1}{2}\right)^5$$

Number of heads (x)	$f(x) = 3,200 \times {}^5C_x \left(\frac{1}{2}\right)^5$
0	$f(0) = 3,200 \times {}^5C_0 \left(\frac{1}{2}\right)^5 = 100$
1	$f(1) = 3,200 \times {}^5C_1 \left(\frac{1}{2}\right)^5 = 500$
2	$f(2) = 3,200 \times {}^5C_2 \left(\frac{1}{2}\right)^5 = 1000$
3	$f(3) = 3,200 \times 5C_3 \left(\frac{1}{2}\right)^5 = 1000$
4	$f(4) = 3,200 \times 5C_4 \left(\frac{1}{2}\right)^5 = 500$
5	$f(5) = 3,200 \times 5C_4 \left(\frac{1}{2}\right)^5 = 100$
Total	$= 3200$

POISSON DISTRIBUTIONS :-

Poisson distribution was derived in 1837 by a French mathematician, Simeon D Poisson (1781 - 1840). In bionomial distribution, the values of p and q and n are given. There is a certainty of the total number of events; in other words, we know the number of times an event does occur and also the times an events does not occur, in binomial distribution. But there are cases where p is very small and n is very large, the calculation involved will be long. Such cases will arise connection with rare events.

The poisson distribution is a discrete distribution with a parameter m. The varies constants are:

1. Mean $= m = p$
2. Standard Deviation $= \sqrt{m}$
3. Skewness given by $= \beta_1 = 1/m$
4. Kurtosis, given by $= \beta_2 = 3 + 1/m$
5. Variance $= m$

Example :-

A book contains 100 mispints distributed randomly throughout its 100 ages. What is the probability that a page obserOved a random contains at least two misprints. Assume poisson distribution.

Solution :-

$$m = \frac{\text{Total number of misprints}}{\text{Total number of pages}} = \frac{100}{100}$$

$$= 1.$$

Probability that a page contains at least two misprints:

$$p(r \geq 2) = 1 - [p(0) + p(1)]$$

$$p(r) = \frac{m^r e^{-m}}{r!}$$

$$P(0) = \frac{1^0 e^{-1}}{0!} = e^{-1} = \frac{1}{e}$$

$$= \frac{1}{2.7183}$$

$$P(1) = \frac{1^1 e^{-1}}{1!} = e^{-1} = 1/e$$

$$= \frac{1}{2.7183}$$

$$P(0) + P(1) = \frac{1}{2.7183} + \frac{1}{2.7183}$$

$$= 0.736$$

$$P(r \geq 2) = 1 - (P(0) + P(1))$$

$$= 1 - 0.736$$

$$= 0.264.$$

Normal Distribution :-

The Binomial distribution and Poisson distribution discussed above are discrete probability distributions. The normal distribution is highly useful in the field to statistics and is an important continuous ~~pa~~ prabability distribution. The graph of this distribution is called normal curve, a bell-shaped curve extending in both the directions, arriving nearer ~~and~~ to the horizantal axis but never touches it.

The normal distribution was first discovered by the English mathematician, De-Moivre (1667 - 1754), in 1733 to solve the problems in game of chances.

Scanned by TapScanner

Example :-

We know that when an unbiased coin is tossed 10 times, the probability of getting x heads is :-

$$P(x) = \frac{n!}{(n-x)!\, x!} (P)^{x} (q)^{10-x}$$

Characteristics of Normal Curve :-

The following points are important properties of normal distribution:

1. The curve is symmetrical. The distribution on the frequencies on either side of the maximum ordinate of the curve is exatly the same. It is a bell-shaped curve. The number of cases above the mean value and below halves would coincide.

2. The value of mean, median and mode will coincide because the distribution is symmetrical and single-peaked.

Mean = Median = Mode

3. μ and σ are the parameters of the normal distribution. For different values of the parameters, we get different normal distribution. The parameters play a central role.

4. It has only mode occuring at μ; it is unimodel.

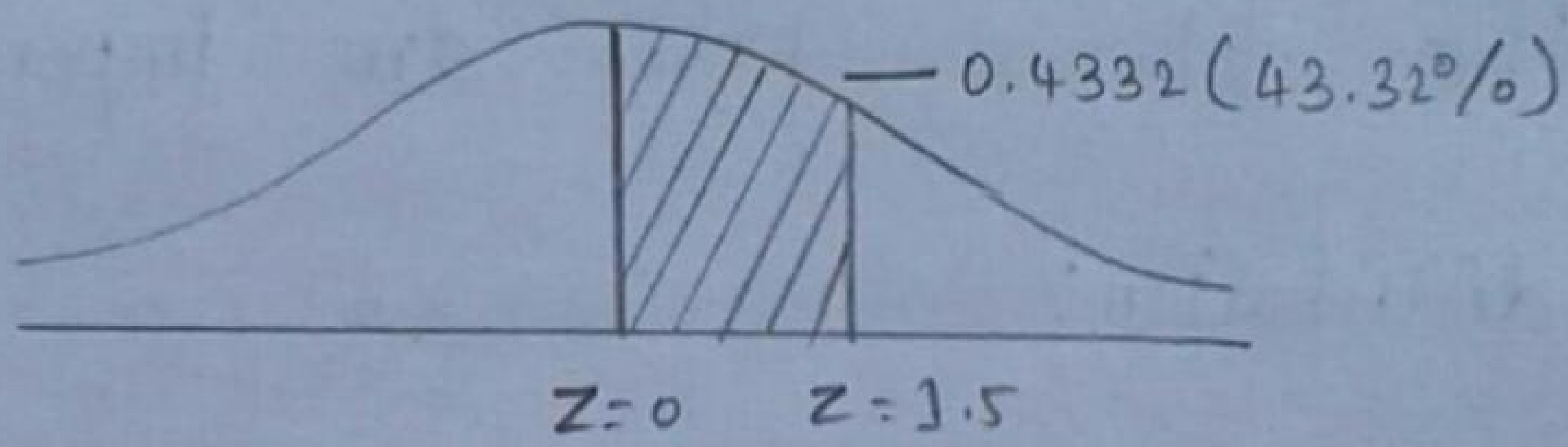
5. Since the distribution is symmetrical, the moment of coefficient of skewness $\beta_1 = 0$, $\beta_2 = 3$

6. The maximum ordinate is at $x = \mu$. Its value is $\frac{1}{\sigma\sqrt{2\pi}}$

Example :-

Find the probability that the standard normal variate lies between 0 and 1.5.

Solution :-



$Z = 0$.

To find the area between $Z = 0$ and $Z = 1.5$, look the area between 0 to 1.5 from the table. It is 0.4332 (shaded area).

Unit - III

Testing of hypothesis - Definition of hypothesis - null hypothesis and alternative hypothesis, standard error, level or significance, critical region, parameters and statistics. Type 1 and Type II errors, one tailed and two tailed tests, Test procedure.

Testing of hypothesis :-

A simple investigation produces results; and with these results, decisions are made on the population. But such decisions involve an element of uncertainty causing wrong decisions. Hypothesis is an assumption which may or may not be true about a population parameter.

The word PARAMETER is used to indicate various statistical measures like, mean, standard deviation, correlation, etc. in the universe. As a against this term STATISTIC refers to the statistical measures relating to the sample. That is parameters are functions of the population values while statistics are functions of the sample observations.

1. Laying Down of Hypothesis :-

To verify our assumption, which is based on sample study, we collect data and find out the different between the sample value and the population value. If there is no different or if the difference between the difference is very small. then our hypothesized value is correct. Generally, two hypothesis must be constructed; and if one hypothesis is correct, the other one is rejected.

a). __Null Hypothesis__:-

It is a very useful tool to test the significance of difference. Any hypothesis concerning a population is called a statistical hypothesis.

i). The average height of the student of a university is 155 cm.

ii). The average daily sales of a firm is Rs. 1500

iii). The average income of a man is a particular village is Rs. 100.

b). __Alternative Hypothesis__:-

Rejection of Ho leads to the acceptance of Alternative hypothesis, which is denoted by H_2.

$$H_0 = \mu = 155$$

$$H_1 = \mu \neq 155 \text{ , ie., } \mu > 155 \text{ or } \mu < 155$$

Scanned by TapScanner

Standard Error :-

If we select a number of independent random samples of a definite size from a given population and calculation som statistics like, mean, standard deviation, etc., from each sample, we shall get a series of values of these statistics.

Example :- In 600 throws of a six faced dice, odd points appeared 360 times. Would you say that the dice is fair at 5% level of significance?

Solution :-

Let us take the hypothesis that the dice is fair.

In a fair dice we would expect 300 odd points in 600 throws.

$$S.E = \sqrt{npq} = \sqrt{600 \times \frac{1}{2} \times \frac{1}{2}} = 12.247$$

$$z = \frac{\text{Difference}}{S.E} = \frac{360-300}{12.247} = 4.89.$$

TEST OF SIGINIFICANCE FOR SMALL SAMPLES

Students t- Distibution :-

When the sample size is 30 or less and the population standard deviation is unknown, we can use the t- distribution.

Assumption for Student t-test :-

1. The parent population from which the samples are drawn is normal.

2. The given sample is random, that is, the given sample is drawn by random sampling.

ONE TAILED ~~et~~ AND TWO TAILED TEST :-

Example :-

Is a correlation coefficient of 0.5 significant if obtained from a random sample of 11 pairs of values from a normal population?

Solution :-

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2}$$

$$t = \frac{0.5}{\sqrt{1-0.5^2}} \times \sqrt{11-2}$$

$$= \frac{0.5}{\sqrt{1-0.5^2}} \times \sqrt{9}$$

$$= \frac{0.5}{0.866} \times 3 = 1.732$$

For $V = 11 - 2 = 9$, $t_{0.005} = 2.26$.

The calculated value of t is less than the table value. Hence, the given correlation ~~coff~~ coefficient is not significant.

Scanned by TapScanner