MAHARAJA INSTITUTE OF TECHNOLOGY MYSORE
(An Autonomous Institution)
Belawadi, Srirangapatna Taluk, Mandya-571477
DEPARTMENT OF MATHEMATICS

# Module-1:Probability Distributions

**1.** Show that the following represents a discrete probability distribution. Find the Mean and Variance.

| X | 10 | 20 | 30 | 40 |
|---|---|---|---|---|
| P(X) | 1/8 | 3/8 | 3/8 | 1/8 |

**2.** The probability distribution of finite random variable is given by the following

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| P(x) | 0 | $k$ | $2k$ | $2k$ | $3k$ | $k2$ | $2k^2$ | $7k^2+k$ |

Find the value of the unknown k. Also find $i)p(3<x\le 6)$ $ii)p(x<6)$ & $iii)p(x\ge 6)$

**3.** Find the value of k such that the following represent a finite probability distribution and find mean and standard deviation. Also find $i)\ p(x\le 1)$ $ii)\ p(x>1)$& $iii)p(-1<x\le 2)$

| x | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| P(x) | k | 2k | 3k | 4k | 3k | 2k | k |

**4.** The finite probability distribution of x is given by the following table,

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| P(x) | k | 3k | 5k | 7k | 9k | 11k | 13k |

For what value of k this represents a valid probability distribution, also find

$i)p(3<x\le 6)$ $ii)p(x<4)$&$iii)p(x\ge 5)$$iv)p(x<3)$

**5.** Show that the give distribution represents a discrete probability distribution, find the mean and variance.

| x | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| P(x) | 0.2 | 0.35 | 0.25 | 0.15 | 0.05 |

**6.** The probability distribution of a finite random variable X is given by the following table. Find the value of k. Also find mean, variance and standard deviation.

| x | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|

| P(x) | 0.1 | k | 0.2 | 2k | 0.3 | k |
|---|---|---|---|---|---|---|

**7.** Find the constant k such that $p(x) = \begin{cases} kx^2, 0 \le x \le 3 \\ 0, \quad otherwise \end{cases}$ Also find $i)\,p\,(x \le 1)\;\; ii)\,p\,(1 \le x \le 2)\;\&\; iii)\,p\,(x > 1)\,iv)\,p\,(x > 2)\,v)\,p\,(x \le 2)$

**8.** The mean and variance of binomial variable x with parameters n and p are 16 and 18. Find $P\,(x \ge 1)\,and\;\; P\,(x > 2)$.

**9.** The probability that a pen manufactured by a company will be defective is 0.1. If 12 such pens are selected, find the probability that, (i) exactly 2 will be defective. (ii) at least 2 will be defective(iii) none will be defective.

**10.** A hospital switchboard receives an average of 4 emergency calls in a 10 minute interval. What is the probability that (i) There are atmost 2 emergency callsin a 10 minute interval (ii) There are exactly 3 emergency call s in a 10 minute interval.

**11.** A dice was thrown 8 times, Find the probability that 3 fall
i) Exactly 2 times (.ii) At least once ( iii) At the most seven times

**12.** In a binomial distribution consisting of 5 independent trials, probabilities of 1 and 2 success are 0.4096 and 0.2048. Find the parameter P of the distribution.

**13.** In a quiz contest of answering i) 'Yes' or 'No' What is the probability of guessing at least 6 answers correctly out of 10 questions asked? ii) Also find the probability of the same if there are four options for a correct answer.

**14.** In 800 families with 5 children each , How many families would be expected to have i) 3 boys (ii) 5 girls (iii) Either 2 or 3 boys (iv).At most 2 girls by assuming probabilities for boys and girls to be equal.

**15.** . An airline knows that 5% of the people making reservations on a certain flight will not turn up.Consequently,their policy is to sell 52 tickets for a flight that can only hold 50 peopleWhat is the probability that there will be a seat for every passenger who turns up?

**16.** The probability that a bomb dropped hits the target is 0.2. Find probability that out of 6 bombs dropped
i) exactly 2 will hit the target ii) at least 3 will hit the target

**17.** If 10% Of the rivets produced by a machine are defective, find the probability that out of 12 randomly chosen rivets: i) Exactly 2 will be defective, ii) At least 2 will be defective, iii) None will be defective.

**18.** The number of telephone lines busy at an instant of time is a binomial variate with probability 0.1 that a line is busy .If 10 lines are chosen at random ,what is the probability that i) no line is busy ii) all lines are busy iii) at least 1 line is busy iv) at most 2 lines are busy.

**19.** The probability that a person aged 60 years will live upto 70 is 0.65. What is the probability that out of 10 persons aged 60 at least 7 of them will live up to 70.

**20.** In a consignment of electric lamps 5% are defective. If a random sample of 8 lamps are inspected what is the probability that one or more lamps are defective?

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

**22.** If the mean and standard deviation of the number of correctly answered questions in a test given to 4096 students are 2.5 and $\sqrt{1.875}$ . Find an estimate of the number of candidates answering correctly.

**23.** The probability that an individual suffers from a bad reaction from a certain injection given for him as a corona suspect,is 0.001. Using Poisson distributiondetermine the probability that out of 2000 individuals, i) Exactly 3 ii)More than 2 will suffer a bad reaction.

**24.** Alpha particles are emitted by a radioactive source at an average of 5 emissions in 20 minutes What is the probability that there will be exactly two emissions i) exactly two emissions. ii) at least two emissions in 20 minutes.

**25.** The number of accidents in a year by taxi drivers in a cityfollows a Poisson distribution with mean 3.Out of 1000 taxi drivers, find approximately the number of taxi drivers with i) no accident in a year ii) More than 3 accidents in a year.

**26.** The number of accidents per day (x) as recorded in a textile industry over a period of 400 days is given.
Fit a Poisson distribution for the data and calculate the theoretical frequencies.

| x | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| P(x) | 173 | 168 | 37 | 18 | 3 | 1 |

**27.** Given 2% of the fuses manufactured by a firm are defective, find by using Poisson distribution the probability that the box containing 200 fuses
i) has no defective fuse
ii)3 or more defective fuses
iii) at least one defective fuse

In a certain factory turning out razor blades there is a small probability of 1/500 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,

i) No defective

ii) One defective iii) Three defective blades in a consignment of 10,000 packets.

Fit a Poisson distribution for the following data and calculate the theoretical frequencies.

| x | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| P(x) | 122 | 60 | 15 | 2 | 1 |

**28.** If x is a exponential variate with mean 5. Evaluatei) p$(0 < x < 1)$

**29.** If x is a exponential variate with mean 3. Find $i$ ) $p$ $(x > 1)$$i$ $i$ ) $p$ $(x < 3)$

**30.** In a certain town the duration of shower has mean 5minutes.What is the probability that shower will last for ( i) 10 minutes or more ( ii) less than 10 minutes (iii) between 10 and 12 minutes.

**31.** A sale per day in a shop is exponentially distributed with the average sale amount of Rs.100 and net profit is 8%. Find the probability that the net profit exceeds Rs.30.
i). In a day ii). In consecutive two days.

**32.** The length of telephone conversation in a booth has been an exponential distribution and found on an average to be 5 minutes . Find the probability that a random call made from this booth . i) Ends in less than 5 minutes ii) Lasts for less than 10 minutes iii) Lasts for 10 mins or more iv) Between 5 and 10 minutes v) Between 10 and 12 minutes

**33.** The length in time (in minutes) that a certain person speaks on the telephone is a random variable with pdf $f(x) = \begin{cases} Ae^{-\frac{1}{5}x} & , \text{for } x > 0 \\ 0 , \text{otherwise} \end{cases}$

Find the value of A and what is the probability that he will talk over the phone,
i) more than 5 minutes ii. less than 10 minutes iii. between 5 and 10 minutes

**34.** The life of a compressor manufactured by a company is known to be 200 months on an average following an exponential distribution. Find the probability that the life of a compressor of that company is i. less than 200 months ii. between 100 months and 25 years.

**35.** The kilo meter (K.M.) run (in 1000 of K.M.s) without any sort of problem in respect of a certain vehicle is a random variable having pdf,

$$f(x) = \begin{cases} \frac{1}{40} e^{-\frac{1}{40}x} & ,\text{for } x > 0 \\ 0 \text{ , otherwise} \end{cases}$$

Find the probability that the vehicle is trouble free i. At least for 25000 KMs

ii. At the most for 25000 KMs

iii. Between 16000 to 32000 KMs.

Hint: Since mean = 40. (in 1000 of K.M.s) ,i) $p(x \leq 25)$ not $p(x \leq 25000)$.

**36.** Find the probabilities of the following:i) $p(z \geq 0.85)$ ii) $p(z \leq -2.43)$ iii) $p(|z| \leq 1.94)$iv) $p(-1.64 \leq z \leq -0.88)$.Given $\varphi(0.85) = 0.3023$,$\varphi(1.64) = 0.4495$,$\varphi(0.88) = 0.3023$,$\varphi(2.43) = 0.4925$, $\varphi(1.94) = 0.4738$.

**37.** If x is a normal variate with 30 and standard deviation 5, find the probabilities that i) $26 \leq x \leq 40$ ii)$x \geq 4$ iii)$|x - 30| > 5$

**38.** In a test on 2000 electrical bulbs, it was found that the life a particular make, was normally distributed with an average life of 2040hours and standard deviation 60 hours Estimate the number of bulbs likely to burn for (a) more than 2150 hours (b) less than 1950 hours (c) more than 1920 hours but less than 2160 hours, given $\varphi$ (1.83) = 04664, $\varphi$ (1.5) = 0.4332,$\varphi$ (2) = 0.4772.

**39.** The marks of 1000 students in an examination follow a normal distribution with mean 70 and standard deviation 5. Find the number of students whose marks will be (i) more than 75 (ii) less than 65 (iii) between 65 and 75.

Given $p(0 < z < 1) = 0.3413$.

**40.** In a normal distribution, 31% of the items are under 45 and 8% of the items are over 64. Find the mean and standard deviation of the distribution, given $\varphi(0.5) = 0.19$, $\varphi(1.4) = 0.42$

**41.** The weekly wages of workers in a company are normally distributed with mean of Rs.700/- & standard deviation of Rs.50. Find probability that weekly wage of a randomly chosen worker is i) between Rs.650 & Rs.750 and ii)more than Rs.750 .

**42.** Suppose that the student IQ scores form a normal distribution with mean 100 and standard deviation 20. Find the percentage of students whose i) score is less than 80 ii) score falls between 90 and 140, iii) score more than 120.

**43.** A sample of 100 dry battery cells tested to find the length of life produced by a company and following results are recorded: mean life =12 hours, standard deviation =3 hours. Assuming data to be normally distributed, find the expected life of a dry cell: i) have more than 15 hours ii) between 10 and 14 hours.

**44.** In an examination 7% of students score less than 35% marks and 89% of students score less than 60% marks .Find the mean and standard deviation if the marks are normally distributed .It is given that $p(0<z<1.2263)=0.39$ and $p(0<z<1.4757)=0.43$

**45.** Steel rods are manufactured to be 3 cms in diameter .But they are acceptable if they are between 2.99cms and 3.01cms .It is observed that 5% are rejected as oversized and 5% are rejected as undersized Assuming that the diameters are normally distributed ,find the standard deviation of the distribution [A(1.645)]=0.45

**46.** Given that the mean heights of the students in a class is 158cms with standard deviation of 20cms. Find how many students height lies in between 150cms and 170cms, if there are 100 students in a class.

## MODULE-2 JOINT PROBABILITY DISTRIBUTION & MARKOV CHAIN

### JOINT PROBABILITY DISTRIBUTION

1. The joint probability distribution of two random variable X and Y is as follows.

| X \ Y | -4 | 2 | 7 |
|---|---|---|---|
| 1 | 1/8 | 1/4 | 1/8 |
| 5 | 1/4 | 1/8 | 1/8 |

Compute the following:
(a) E(X) and E(Y) (b) E (XY) (c) $\sigma_x$ and $\sigma_y$
(d) COV(X,Y) (e) $\rho(X,Y)$

2. The joint probability distribution of two random variable X and Y is as follows.

| X \ Y | -2 | -1 | 4 | 5 |
|---|---|---|---|---|
| 1 | 0.1 | 0.2 | 0 | 0.3 |
| 2 | 0.2 | 0.1 | 0.1 | 0 |

Determine the marginal probability distribution of X and Y.

Also compute the following. (a) E(X) and E(Y) (b) E (XY) (c) $\sigma_x$ and $\sigma_y$
(d) COV(X,Y) (e) $\rho(X,Y)$
Further verify that X and Y are dependent random variables. Also find $P(X+Y>0)$

3. Suppose X and Y are independent random variables with the following respective distribution, find the joint distribution of X and Y. Also verify that COV(X,Y) = 0

| x | 1 | 2 |
|---|---|---|
| f(x) | 0.7 | 0.3 |

| y | -2 | 5 | 8 |
|---|---|---|---|
| g(y) | 0.3 | 0.5 | 0.2 |

4. X and Y are independent random variables. X takes values 2, 5, 7 with probability 1/2 ,1/4 , 1/4 respectively. Y takes values 3,4,5 with probability 1/3 , 1/3 , 1/3.
   (i) Find the joint probability distribution of X and Y.
   (ii) Shoe that COV(X,Y) = 0
   (iii) Find the probability distribution of Z=X+Y

5. Given the following joint probability distribution of random variable X and Y, find the corresponding marginal probability distribution. Also compute the COV(X,Y) and ρ(X,Y)

| X \ Y | 1 | 3 | 9 |
|---|---|---|---|
| 2 | 1/8 | 1/24 | 1/12 |
| 4 | 1/4 | 1/4 | 0 |
| 6 | 1/8 | 1/24 | 1/12 |

6. The joint probability distribution of two random variable X and Y is given by f( x , y) = k(2x+y) where x and y are integers such that $0 \leq x \leq 2$ , $0 \leq y \leq 3$ .
   (i) Find the value of the constant k
   (ii) Find the marginal probability distribution of X and Y.
   (iii) Show that the random variables X and Y are dependent.
   (iv) E(X) , E(Y) , E (XY) , $\sigma_x$ and $\sigma_y$
7. A fair coin is tossed thrice. The random variables X and Y are defined as follows.
   X = 0 or 1 according as head or tail occurs on the first toss
   Y = Number of heads.
   (i) Determine the distribution of two random variable X and Y
   (ii) Determine joint probability distribution of two random variable X and Y
   (iii) Compute the COV(X,Y) and ρ(X,Y)
   (iv) P (X =1 , Y = 2) P($X \leq 1$ , $Y \leq 2$)
8. Random variables X and Y have the following joint probability distribution

| X \ Y | -3 | 2 | 4 |
|---|---|---|---|
| 1 | 0.1 | 0.2 | 0.2 |
| 2 | 0.3 | 0.1 | 0.1 |

   (i) Determine the marginal probability distribution of X and Y.
   (ii) Compute the COV(X,Y) and ρ(X,Y)
   (ii) Are the variables X , Y independent random variables ?

9. Let X be a random variable with the following distribution and Y is defined to be $X^2$

| x | -2 | -1 | 1 | 2 |
|---|---|---|---|---|
| f(x) | 1/4 | 1/4 | 1/4 | 1/4 |

Determine

(i)the distribution g of Y (ii) joint probability distribution X and Y

(iii) E(X) , E(Y) , E (XY) , COV(X,Y) and ρ(X,Y)

10. A coin is tossed thrice. The random variables X and Y are defined as follows.

    X = 0 or 1 according as tail or head occurs on the first toss

    Y = Number of tails.

    (i)Determine the distribution of two random variable X and Y

    (ii) Determine joint probability distribution of two random variable X and Y

    (iii)Find E(X+Y) and E(XY)

## MARKOV CHAIN

1. If $A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$ is a stochastic matrix and $v = [v_1 \quad v_2]$ is a probability vector, show that $vA$ is also a probability vector.
2. Find the unique probability vector of the regular stochastic matrix $A = \begin{bmatrix} 3/_4 & 1/_4 \\ 1/_2 & 1/_2 \end{bmatrix}$
3. Find the unique probability vector of the regular stochastic matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 1/6 & 1/2 & 1/3 \\ 0 & 2/3 & 1/3 \end{bmatrix}$
4. Find the unique probability vector of the regular stochastic matrix $A = \begin{bmatrix} 0 & 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}$
5. Show that $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix}$ is regular stochastic matrix. Also find the associated unique probability vector.
6. The transition matrix $P$ of a Markov chain is given by $\begin{bmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{bmatrix}$ with the initial probability distribution $P^{(0)} = \left(\frac{1}{4}, \frac{3}{4}\right)$. Define and find the following.

   $(i)\ \ P_{21}^{(2)} \quad (ii)\ \ P_{12}^{(2)}\ \ (iii)\ P^{(2)} \quad (iv)\ \ P_1^{(2)} \quad (v) the\ vector\ P^{(0)}\ \ P^n\ \ approches.$

   $(vi)\ the\ matrix\ P^n\ \ approches.$

7. The transition probability matrix of Markov chain is given by $P = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$ and the initial probability distribution is $P^{(0)} = (\ 1/2\ , 1/2, 0)$ . Find $P_{13}^{(2)}$ , $P_{23}^{(2)}$, $P^{(2)}$ $and$ $P_1^{(2)}$.

8. Prove that the Markov chain whose t.p.m is $P = \begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$ is irreducible. Find the corresponding stationary probability vector.

9. Prove that the Markov chain whose t.p.m is $P = \frac{1}{10}\begin{bmatrix} 6 & 2 & 2 \\ 1 & 8 & 1 \\ 6 & 0 & 4 \end{bmatrix}$ is irreducible. Find the corresponding stationary probability vector.

10. Show that $A = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix}$ is regular stochastic matrix. Also find the associated unique probability vector.

11. A software engineer goes to his office everyday by motorbike or by car. He never goes by bike on two consecutive days, but if he goes by car on a day then he is equally likely to go by car or by bike the next day. Find the t .p .m of the Markov chain. If car is used on the first day of the week find the probability that after 4 days (a) bike is used (b) car is used.

12. A salesman's territory consists of 3 cities A,B, C. He never sells in the same city for 2 consecutive days. If he sells in city A, then the next day he sells in city B. However if he sells in either B or C , then the next day he is twice as likely to sell in city A as in the other city. In the long run how often does he sell in each of the cities?

13. A habitual gambler is a member of two clubs A and B. he visits either of the clubs every day for playing cards. He never visits club A on two consecutive days. But , if visits club B on a particular day, then the next day he is as likely to visit club B or club A. Find the t . p. m of this Markov chain. Also, (a) show that the matrix is a regular stochastic matrix and find the unique probability vector. (b) if the person had visited club B on Monday , find the probability that he visits club A on Thursday.

14. A student's study habits are as follows. If he studies one night, he is 70% sure not to study the next night. On the other hand if he does not study one night, he is 60% sure not to study the next night. In the long run how often does he study?

15. A man's smoking habits are as follows. If he smokes filter cigarettes one week, he switches to non-filter cigarettes the next week with probability 0.2. On the other hand

, if he smokes non filter cigarettes one week there is a probability of 0.7 that he will smoke non filter cigarettes the next week as well. In the long run how often does he smoke filter cigarettes.

16. Each year a man trades his car for a new car in 3 brands of the popular company Maruti Udyog limited. If he has a Standard he trades it for Zen. If he has a Zen he trades it for Esteem. If he has a Esteem he is just as likely to trade it for a new Esteem or for a Zen or a Standard one. In 1996 he brought his first car which was Esteem. Find the probability that he has (a) 1998 Esteem (b) 1998 Standard (c) 1999 Zen (d) 1999 Esteem.

17. Three boys A ,B ,C, are throwing ball to each other. A always throw the ball to B and B always throw the ball to C. C is just as likely to throw the ball to B as to A. If C was the first person to throw the ball find the probabilities that after three throws (a) A has ball (b) B has ball
(c) C has ball.

18. Two boys $B_1$ , $B_2$ and two girls $G_1$ , $G_2$ are throwing ball from one to the other. Each boy throws the ball to the other boy with probability 1/2 and to each girl with probability 1/4. On the other hand each girl throws the ball to each boy with probability 1/2 and never to the other girl. In the long run how often does each receive the ball.

19. A gambler's luck follows a pattern. If he wins a game, the probability of winning the next game is 0.6. However if he loses a game, the probability of losing the next game is 0.7. There is an even chance of gambler winning the first game. If so (a) what is the probability of he winning the second game? (b) what is the probability of he winning the third game? (b) In the long run, how often he will win?

## Module-3-Statistical Inference 1

1. Define sampling distribution?
2. Define standard error.
3. A coin was tossed 400 times and the head turned up 216 times. Test the hypothesis that the coin is unbiased at 5% level of significance.
4. A die was thrown 9000 times and a throw of 5 or 6 was obtained 3240 times. On the assumption of random throwing do the data indicate an unbiased die?
5. In a locally containing 18000 families a sample of 840 families was selected at random. Of these 840 families were found to have a monthly income of 250 Rs or less. It is desired to estimate how many out of 18,000 families have a monthly income of 250 Rs or less. Within what limits would you place your estimate?
6. In a city A 20% of random sample of 900 school boys had a certain slight physical defect. In another city B 18.5% of a random sample of 1600 school boys had the same defect. Is the difference between the proportions significant?
7. In 2 large populations there are 30% and 25% respectively of fair-haired people. Is this difference likely to be hidden in samples of 1200 and 900 respectively from the two populations?

8. In a sample of 1000 people in Karnataka 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance?
9. 20 people were attacked by a disease and only 18 survived. Will you reject the hypothesis that the survival rate. If attacked by this disease is 85% in favour of the hypothesis that it is more at 5% level?
10. In a random sample of 400 persons from a large population 120 are female. Can it be said that males and females are in the ratio 5:3 in the population? Use 1% level of population.
11. In a big city 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers?
12. Can it be concluded that the average life span of Indian is more than 70 years. If random sample of 100 Indians has an average life span of 71.8 years with SD of 7.8 years. (Assumes 5% level of significance)
13. A die is thrown 600 times and the digits 2 or 4 is considered as success. Digit 2 or 4 are obtained for 212 times. Is a die unbiased?
14. A coin is tossed 1000 times and heads is received for 280 times. Is the coin unbiased? (Use 1% level of significance)
15. In a sample of 500 people from a state 280 takes tea and rest take coffee. Can we assume that tea and coffee are equally popular in the state at 5% level of significance?
16. A machine produces 16 imperfect articles in a sample of 500. After machine is overhauled, it produces 3 imperfect articles in a batch of 100. Has the machine been improved?
17. A die is tossed 960 times and it falls with 5 upwards 184 times. Is the die biased?
18. In a sample of 600 men from a certain city, 450 are found smokers. In another sample of 900 men from another city, 450 are smokers. Do the data indicate that the cities are significantly different with respect to the habit of smoking among men?
19. In a group of 50 first cousins there were found to be 27 males and 23 females. Ascertain if the observed proportions are inconsistent with the hypothesis that the gender should be in equal proportions.
20. Ball are drawn from a bag containing equal number of black and white balls, each ball being replaced before drawing another. In 2250 drawing 1018 black and 1232 white balls have been drawn. Do you suspect some bias on the part of the drawer?

## Module-4 (Statistical Inference-2)

1) Define Sampling variable , central limit theorem , confidence limit for unknown mean.
2) A Population consists of 5,10,14,18,13,24 . consider all Possible samples of size two which can be drawn without replacement from the population Find

   a) Mean of the Population b) S.D of the Population

   c) Mean of sampling distribution of means,

   d) S.D of sampling distribution of means

3) A random sample of size 100 is taken from an infinite population having $\mu = 76$ varince $\sigma^2 = 256$, what is the probability that $\bar{x}$ will be between 75 & 78.

4) A normal population has mean of 0.1 & S.D of 2% Find the probability that mean of a sample size 900 will be negative.

5) A sample size of 64 and mean 60 was taken from population whose S.D is 10. Construct 95 % confidence interval for the mean.

6) A certain stimulus administered to each of 12 patients resulted in the following increases of blood pressure: 5, 2, 8, -1, 3, 0, -2, 1, 5, 0, 4, 6. Can it be concluded that the stimulus will in general be accompanied by an increase in blood pressure.

7) The nine items of a sample have the following values: 45, 47, 50, 52, 48, 47, 49, 53 and 51. Does the mean of these differ significantly from the assumed mean of 47.5?

8) Ten individuals are chosen at random from a population and their heights in inches are found to be 63, 63, 66, 67, 68, 69, 70, 70, 71, 71 . Test the hypothesis that the mean height of the universe is 66 inches. (For d.f. 9, $t_{0.05} = 2.262$ )

9) A machinist is making engine parts with axle diameter of 0.7 inch. A random sample of 10 parts shows mean diameter 0.742 inch with a S.D 0.04 inch. On the basis of this sample, would you say that the work is inferior?

10) Eleven school boys were given a test in drawing. Further they were given a month's tuition and a second test of equal difficulty was held at the end of it. Do the marks give the evidence that students have benefitted by extra coaching? For d.f. $\gamma = 10$, $t_{0.05} = 2.228$.

11) In experiments on pea breeding, the following frequencies of seeds were obtained.

| Round and yellow | Wrinkled and yellow | Round and green | Wrinkled and green | Total |
|---|---|---|---|---|
| 315 | 101 | 108 | 32 | 556 |

Theory predicts that the frequencies should be in proportions $9:3:3:1$. Examine the correspondence between theory and experiment.

12) A set of five similar coins is tossed 320 times and the result is

| No. of heads | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f$ | 6 | 27 | 72 | 112 | 71 | 32 |

Test the hypothesis that the data follow a binomial distribution.

13) A sample of height of 6400 soldiers has a mean of 67.85 inches and S.D. 2.56 inches while a sample of heights of 1600 sailors has a mean of 68.55 inches and S.D. 2.52 inches. Do the data indicate that sailors are on the average taller than the soldiers?

14) From the data given below about the treatment of 250 patients suffering from a disease, state whether new treatment is superior to the conventional test.

| Data | Number of patients | | |
|---|---|---|---|
| | Favourable | Not favorable | Total |
| New one | 140 | 30 | 170 |
| Conventional | 30 | 20 | 80 |
| Total | 200 | 50 | 280 |

15) The manufacturer of a certain make of electric bulbs claims that his bulbs have a mean life of 25 months with a standard deviation of 5 months. Random samples of 6 such bulbs

have the following values: Life of bulbs in months: 24, 20, 30, 20, 20, and 18. Can you regard the producer‟s claim to valid at 1% level of significance?

16) The life time of electric bulbs for a random sample of 10 from a large consignment gave the following data: 4.2, 4.6, 3.9, 4.1, 5.2, 3.8, 3.9, 4.3, 4.4, 5.6 (in 000 hours). Can we accept the hypothesis that the average life time of bulbs is 4,000 hours?

17) Two horses A and B were tested according to the time (in seconds) to run a particular race with the following results.

Horse A: 28 30 32 33 33 29 34

Horse B: 29 30 30 24 27 29 Test whether you can discriminate between the two horses.

18) 11 students were given a test in statistics, they were provided additional coaching and then a second test of equal difficulty was held at the end of the coaching. Marks scored by then in the two tests were given

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Test 1 | 23 | 20 | 19 | 21 | 18 | 20 | 18 | 17 | 23 | 16 | 19 |
| Test 2 | 24 | 19 | 22 | 18 | 20 | 22 | 20 | 20 | 23 | 20 | 17 |

Do the marks given evidence that the students have benefited by extra coaching? Given $t_{0.05}$ (10)=2.228 test the hypothesis at 5% level of significance.

19) Genetic theory state that children having parent of blood type M and the other of blood type N, will always be one of three types M, MN, N and that the proportion of these types on an average be 1:2:1. the report states that one of the 30 children having one M parent and one N parent, 30% are found to be of type M, 45% are found to be of type MN, and the remainder of type N, test the theory by chi square test.

20)The blood pressure of 5 woman before and after intake of a certain drug are given below

| | | | | | |
|---|---|---|---|---|---|
| Before | 110 | 120 | 123 | 132 | 125 |
| After | 120 | 118 | 125 | 136 | 121 |

Test whether there is significant change in blood pressure at 1% level of significance ?

21) The Nichols contents in milligrams in sample of Tabaco were found to be as follows

| | | | | | | |
|---|---|---|---|---|---|---|
| Sample A | 24 | 27 | 26 | 21 | 25 | - |
| Sample B | 27 | 30 | 28 | 31 | 22 | 36 |

22) Pumpleins were grown under two experimental conduction. Two random samples of 11 & 10 , Show the sample S.D of their weights as 0.8 and 0.5 respectively . Assume that the weight distribution as normal , Test the hyp is that the true variance are equal.

## **Module 5-Design of Experiments & ANOVA**

1. What are the principles of experimentation in ANOVA.
2. What is ANOVA?
3. 3 Types of fertilizers A, B, C are used on 12 plots of cultivated land in order test their effect on productivity. Fertilizers are applied to each plot randomly. Details of

productivity is given below. Test the hypothesis that productivity of all 3 fertilizers is same.

| Plot | Production (metric tons) using fertilizers A, B, C | | |
|---|---|---|---|
| | A | B | C |
| 1 | 3 | 10 | 5 |
| 2 | 4 | 7 | 4 |
| 3 | 6 | 8 | 5 |
| 4 | 4 | -6 | 5 |

4. A completely randomised design experiment with 10 plots and 3 treatment gave the following result.

| Plot number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | A | B | C | A | C | C | A | B | A | B |
| Yield | 5 | 4 | 3 | 7 | 5 | 1 | 3 | 4 | 1 | 7 |

Analyse the result for treatment.

5. 3 varieties of a crop are tested in a randomised block design with 4 replications, the layout being as given below. the yields are given in kilograms analyse for significance.

| | | | |
|---|---|---|---|
| C48 | A51 | B52 | A49 |
| A47 | B49 | C52 | C51 |
| B49 | C53 | A49 | B50 |

6. 5 breeds of cattle $B_{1,}$ $B_{2,}$ $B_{3,}$ $B_{4,}$ $B_5$ were fed on 4 different rations $R_{1,}$ $R_{2,}$ $R_{3,}$ $R_4$ gain in weight in kg. over a given period were recorded and given below.

| | $B_1$ | $B_2$ | $B_3$ | $B_4$ | $B_5$ |
|---|---|---|---|---|---|
| $R_1$ | 1.9 | 2.2 | 2.6 | 1.8 | 2.1 |
| $R_2$ | 2.5 | 1.9 | 2.3 | 2.6 | 2.2 |
| $R_3$ | 1.7 | 1.9 | 2.2 | 2.0 | 2.1 |
| $R_4$ | 2.1 | 1.8 | 2.5 | 2.3 | 2.4 |

Is there a significant difference between a. breeds & b. rations

7. In order to determine whether there is significant difference in the durability of 3 makes of computers samples of size 5 are selected from each make & the frequency of repair during the 1st year of purchase is observed. The result are as follows.

| Makes | | |
|---|---|---|
| A | B | C |
| 5 | 8 | 7 |
| 6 | 10 | 3 |
| 8 | 11 | 5 |
| 9 | 12 | 4 |
| 7 | 4 | 1 |

8. What are the basic principles of ANOVA.
9. To access the significance of possible variation in performance in a certain test between the school of A city. A common test was given to a number of students taken at random from the 12th class of the 3school concerned. The results given below. Make the ANOVA for the given data.

| A | B | C |
|---|---|---|
| 2 | 3 | 4 |
| 4 | 5 | 6 |
| 6 | 7 | 8 |

10. To study the performance of 3 detergents and 3 different water temperature the following whiteness readings were obtained with specially designed equipment.

| Water temperature | A | B | C |
|---|---|---|---|
| Cold water | 47 | 45 | 50 |
| Warm water | 39 | 42 | 52 |
| Hot water | 44 | 36 | 48 |

11. Below are given the yield (in kg) per acre for 5 trial plots of 4 varieties of treatment

| Plot number | Treatment | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| 1 | 42 | 48 | 68 | 80 |
| 2 | 50 | 66 | 52 | 94 |
| 3 | 62 | 68 | 76 | 78 |

| | | | | |
|---|---|---|---|---|
| 4 | 34 | 78 | 64 | 82 |
| 5 | 52 | 70 | 70 | 66 |

Carryout an analysis of variance & state your conclusions

12. To test the significance of the variation of the retail prices of a commodity in 3 cities 4 shops were choosen at random in each city & prices observed in rupees were as follows.

| | Shops | | | |
|---|---|---|---|---|
| Bombay | 16 | 8 | 12 | 14 |
| Culcutta | 14 | 10 | 10 | 16 |
| Delhi | 4 | 10 | 8 | 8 |

To the data indicates that the prices in 3 cities are significantly different or not at 5%.

13. 3 varieties of coal were analysed by 4 chemists and the ash contents in the varieties was found as

| Varieties | Chemists | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| A | 8 | 5 | 5 | 7 |
| B | 7 | 6 | 4 | 4 |
| C | 3 | 6 | 5 | 4 |

Discuss the significance of the difference between

a. Chemists
b. Varieties of coal in respect of ash content.

14. Perform a 2way ANOVA on the data given below

| Teachers | Students | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| A | 30 | 24 | 33 | 36 | 27 |
| B | 26 | 29 | 24 | 31 | 35 |
| C | 38 | 28 | 35 | 30 | 35 |

a. Shift the given origin to 30. Perform the ANOVA for the transformed data.

b. How do the results compare with those obtained for the original data?

15. Given the following data test the hypothesis.
    $H_0$: all the means are equal
    $N_1=4$, $\bar{x}_1=27$, $S_1=4$, $N_2=7$, $\bar{x}_2=25$, $S_2=9$, $N_3=5$, $\bar{x}_3=28$, $S_3=5$
    a. Is it one way or two way ANOVA?
    b. Construct ANOVA table.

16. Set up a two way ANOVA table for the following per hectare yield for 4 varieties of yield of wheat on 3 plots.

| Plot of land | Yield | | | |
|---|---|---|---|---|
| | A | B | C | D |
| 1 | 3 | 4 | 6 | 6 |
| 2 | 6 | 4 | 5 | 3 |
| 3 | 6 | 6 | 4 | 7 |

17. Construct the ANOVA table for the following information

| Drivers | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| a | 18A | 21B | 25C | 11D |
| B | 22B | 12C | 15D | 19A |
| C | 15C | 20D | 23A | 24B |
| d | 22D | 21A | 10B | 17C |

18. Analyse the variance in the following Latin square of yields (in kgs) of paddy where A, B, C, D denote the different methods of cultivation

| | | | |
|---|---|---|---|
| D122 | A121 | C123 | B122 |
| B124 | C123 | A122 | D125 |
| A120 | B119 | D120 | C121 |
| C122 | D123 | B121 | A122 |

Examine whether the different methods of cultivation have given significantly different yields.

19. Explain randomised block design briefly.
20. Explain Latin square design briefly.