VALLIAMMAI ENGINEERING COLLEGE

DEPARTMENT OF MATHEMATICS

MA7155 – APPLIED PROBABILITY AND STATISTICS

(UNIVERSITY QUESTIONS & QUESTION BANK)

(PREPARED BY V.VIJAYALAKSHMI)

UNIT- I ONE DIMENSIONAL RANDOM VARIABLES

PART – A

1. **Define random variable.**
2. **A continuous random variable X has a pdf $f(x) = 3x^2$, $0 \le x \le 1$, Find a and b such that $P(x \le a) = P(x > a)$ and $P(x > b) = 0.05$**
3. **The first 4 moments about 3 are 1 and 8. Find the mean and variance**
4. **Define moment generating function of a random variable X.**
5. **If X is a Poisson variate such that P (X =2) =9 P (X =4) +90 P (X = 6) .Find the variance of X.**
6. **If X is a continuous RV with p.d.f. f(x) =2x ,0 <x< 1 , then find the pdf of the RV $Y = X^3$**
7. **If the mean of a Poisson varaite is 2,then what is the standard deviation?**
8. **If X and Y are independent RVs with varainaces 2 and 3 .Find thevariance of 3X + 4Y.**
9. **The first four moments of a distribution about 4 are 1 ,4,10 and 45 respectively. Show that the mean is 5 and variance is 3.**
10. $If\ f(x) = \begin{cases} kx^2 & ,0 < x < 3 \\ 0, otherwise \end{cases}$ **Find the value of X ,then find the value of K**
11. **The random variable X has a Binomial distribution with parameters n =20 ,p =0.4 .Determine P (X = 3).**
12. **If the probability is 0.05 that a certain kind of measuring device will show excessive drift , what is the probability that the sixth measuring device tested will be the first to show excessive drift?**
13. **If a RV has the probability density $f(x) = \begin{cases} 2e^{-2x} & for\ x > 0 \\ 0, for\ x \le 0 \end{cases}$, find the probabilities that will take a value between 1 and 3.**

14. A RV X has the p.d.f. $f(x) = \begin{cases} xe^{-x} \; for \; x > 0 \\ 0, for \; x \leq 0 \end{cases}$ .Find the CDF of X

15.If the RV X takes the values 1 ,2,3 and 4 such that 2P(X = 1)=3P(X =2) =P (X =3) = 5 P (X =4) find the probability distribution .

16.State the memoryless property of an exponential distribution.

17. If a RV has the probability density $f(x) = \begin{cases} 2e^{-2x} \; for \; x > 0 \\ 0, for \; x \leq 0 \end{cases}$, find the mean and variance of the RV X.

18. Obtain the moment generating function of Geometric distribution.

19. Given that the p.d.f of a random variable X is f(x) = kx, 0 < x < 1 , find k.

20. Find the Binomial distribution for which the mean is 4 and variance is 3.

PART – B

1.Buses arrive at a specific stop at 15 minutes intervals starting at 7a.m.If a passenger arrives at a random time that is uniformly distributed between 7 and 7.30 am , find the probability that he waits 1) less than 5 minutes for a bus and 2) atleast 12 minutes for a bus .

2. A manufacturer of certain product knows that 5 % of his product is defective .If he sells his product in boxes of 100 and guarantees hat not more than 10 will be defective ,what is the approximate probability that a box will fail to meet the guaranteed quality?

3. If X is a discrete random variable with probability function p(x) = $\frac{1}{K^x}$, x =1,2 ... (K constant) then find the moment generating function ,mean and variance..

4. In a company ,5% defective components are produced .What is the probability that at least 5 components are to be examined in order to get 3 defectives?

5. 12. Let the random variable X has the p.d.f. $f(x) = \begin{cases} \frac{1}{2} e^{-\frac{x}{2}} , x > 0 \\ 0 \, . \, otherwise \end{cases}$

Find the mean and variance.

6. In a normal distribution, 31 % of the items are under 45 and 8% are over 64.Find the mean and variance of the distribution.

7. If X is Uniformly distributed over (0 ,10) ,find the probability that

(i) X < 2 (ii) X > 8 (iii) 3 <X <9?

8.A discrete RV X has the probability function given below

X :	0	1	2	3	4	5	6	7

P(x) : 0 a 2a 2a 3a a^2 2 a^2 $7a^2 + a$

Find (i) Value of a (ii) p (X <6) , P (X $\geq$ 6) , P (0 < X < 4) (iii) Distribution function.

9. The slum clearance authorities in a city installed 2000 electric lamps in a newly constructed township. If the lamps have an average life of 1000 burning hours with a standard deviation of 200 hours , 1) what number of lamps might be expected to fail in the first 700 burning hours? 2) After what period of burning hours would you expect 10 percent of the lamps would have been failed ?

(Assume that the life of the lamps follows a normal law)

10. The daily consumption of milk in a city ,in excess of 20,000 gallons ,in approximately distributed as a Gamma variate with the parameters k = 2 and $\lambda = \frac{1}{10{,}000}$.The city has a daily stock of 30,000 gallons. What is the probability that the stock is insufficient on a particular day?

11. The p.d.f of a random variable X is given by f(x) = kx (2-x), 0 ≤ x≤ 1, Find k , mean , variance rth moment.

12. The weekly wages of 1000 workmen are normally distributed around a mean of Rs. 70 and with standard deviation of Rs. 5. Estimate the number of workers whose weekly wages will be between Rs.69 and Rs. 72, less than Rs.69, more than Rs.72.

13. A random variable X has a uniform distribution over (-3,3) compute P(X<2), P([x]<2) , P([x-2]≤ 2).

14. A random variable X has the following probability distribution

X	-2	-1	0	1	2	3
P(X)	0.1	k	0.2	2k	0.3	3k

Find i) k ii) P(X<2) iii) P(-2 < X < 2) iv) P(X > 1).

15. State and Prove memoryless property of Exponential distribution.

UNIT – II TWO DIMENSIONAL RANDOM VARIABLES

PART –A

1. **The joint pdf of a bivariate RV (X,Y) is given by**

$$f(x,y) = \begin{cases} 4xy & 0 < x < 1,\ \ 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$$

Find *P(X+Y<1)*.

2. **Consider the two – dimensional density function.**

$$f(x,y) = 2 \ \text{ for } \begin{cases} 0 < x < 1 \\ 0 < y < x \end{cases}$$

$f(x,y) = 0$, outside. **Find the marginal density functions.**

3. **If the joint pdf of (X,Y) is** $f(x,y) = 6e^{-2x-3y},\ x \geq 0, y \geq 0,$ **find the marginal density of *X* and conditional density of *Y* given *X*.**

4. **The joint pdf of (X,Y) is given by** $f(x,y) = e^{-(x+y)}, 0 \leq x, y < \infty.$ **Are X and Y independent? Why?**

5. **Given the joint density function of X and Y as**

$$f(x,y) = \frac{1}{2}e^{-y}, 0 < x < 2, y > 0$$
$$= 0, \quad \text{elsewhere}$$

find the distribution function of (X+Y).

6. **The following table gives the joint probability distribution of X and Y. Find the a) marginal density function of X. b) marginal density function of Y.**

Y \ X	1	2	3
1	0.1	0.1	0.2
2	0.2	0.3	0.1

7. **The joint probability mass function of (X,Y) is given by P(x,y) = K (2x+3y), *x* = 0, 1,2, y = 1,2,3. Find the marginal probability distribution of X : {i, P_i}**

X \ Y	1	2	3
0	3K	6K	9K
1	5K	8K	11K
2	7K	10K	13K

8. **Find the value of *k*, if** $f(x,y) = k(1-x)(1-y),$ **for 0< x, y<1, is to be a joint density function.**

9. **Let *X* be a random variable with pdf** $f(x) = \frac{1}{2}, -1 \leq x \leq 1,$ **and let $Y=X^2$. Prove that correlation co-efficient between *X* and *Y* is zero.**

10. **Prove that the correlation co-efficient lies between +1 and -1.**

11. **Find the mean values of the variables *X* and *Y* and correlation co-efficient from the following regression equations:**

2Y – X – 50 = 0

3Y – 2X – 10 = 0

12. **The correlation co-efficient between two random variables *X* and *Y* is *r = 0.6.* If** $\sigma_X = 1.5,\ \sigma_Y = 2,\quad \overline{X} = 10$ and $\overline{Y} = 20$ **, find the regression of (i) Y on X and (ii) X and Y.**

13. **The following results were worked out from scores in Maths (X) and Statistics (Y) of students in an examination:**

	X	*Y*
Mean	39.5	47.5
Standard deviation	10.8	17.8

Karl Pearson's correlation co-efficient = +0.42

Find both the regression lines. Use these regressions and estimate the value of Y for X .

14. **The co-efficient of correlation between *x* and *y* is 0.48. Their covariance is 36. The variance of *x* is 16. Find the standard deviation of *y*.**

15. **If *x, y* denote the deviations of the variates from the arithmetic means and if**

$$r = 0.5,\ \sum xy = 120,\ \sigma_y = 8,\ \sum x^2 = 90$$ **, find *n*, the number of items.**

16. **If the random variable *X* is uniformly distributed in (0,1) and** $Y = X^2$**, find (i) V(Y) (ii)** r_{XY}**.**

17. **If the random variable X is uniformly distributed over (-1,1), find the density function of** $Y = \sin\left(\frac{\pi x}{2}\right)$

18. **The bivariate random variable X and has pdf f(x,y) = kxy, for 0 < x < 4, 1 < y < 5. Find k**

19. **Write any two properties of correlation coefficient.**

20. **The two regression lines are 4x – 3y +33 = 0, 20x – 9 y = 107, var(x) = 25, Find the mean of x and y.**

PART – B

1. **The two dimensional RV(X,Y) has the joint density**

$$f(x,y) = 8xy,\ 0 < x < y < 1$$
$$= 0, \text{ otherwise}$$

(i) **Find** $P(X < 1/2 \cap Y < 1/4)$,

(ii) **Find the marginal and conditional distributions, and**

(iii) **Are X and Y independent? Give reasons for your answer.**

2. ***Let the RV X has the marginal density function*** $g(x) = 1, -\frac{1}{2} < x < \frac{1}{2}$ ***and let the conditional density of Y be***

$$h(y/x) = 1, \quad x < y < x+1, \quad -\tfrac{1}{2} < x < 0$$
$$= 1, \quad -x < y < 1-x, \quad 0 < x < \tfrac{1}{2}$$

Show that the variables X and Y are uncorrelated.

3. **If the joint pdf of a two – dimensional RV (X,Y) is given by**

$$f(x,y) = K(6-x-y);\ 0 < x < 2,\ \ 2 < y < 4$$
$$= 0, \text{ elsewhere}$$

find (i) the value of *K*, (ii) $P(X < 1, Y < 3)$ **(iii)** $P(X + Y < 3)$
(iv) $P(X < 1 / Y < 3)$

4. **Determine the value of C that makes the function *F*(*x* , *y*) = C (*x* + *y*) a joint probability density function over the range 0 < *x* < 3 and *x* < *y* < *x* + 2. Also determine the following.**

   i) ***P* (*X* < 1, *Y*< 2)**

   ii) ***P* (*Y* > 2)**

   iii) ***E* [*X*]**

5. **If the joint pdf of a two – dimensional RV (X,Y) is given by**

$$f(x,y) = x^2 + \frac{xy}{3}\ ; \quad 0 < x < 1, 0 < y < 2$$
$$= 0, \text{ elsewhere}$$

find (i) P(X > ½), (ii) P(Y<X) and (iii) P(Y< ½ / X < ½).

6. **Let X_1 and X_2 be two RVs with joint pdf given by** $f(x_1, x_2) = e^{-(x_1 + x_2)}$; $x_1, x_2 \geq 0$ **otherwise. Find the marginal densities of X_1 and X_2. Are they independent? Also find** $P[X_1 \leq 1, X_2 \leq 1]$ **and** $P(X_1 + X_2 \leq 1)$.

7. **The joint distribution of X_1 and X_2 is given by** $f(x_1, x_2) = \frac{x_1 + x_2}{21}, x_1 = 1, 2$ **and 3 ;** x_2 **= 1 and 2. Find the marginal distributions of X_1 and X_2.**

8. **If X and Y are independent random variables with density function** $f_X(x) = 1$ **in** $1 \leq x \leq 2$ $\quad f_Y(y) = \frac{y}{6}$ **in** $2 \leq y \leq 4$, **find the density function of Z = XY.**

9. **If X, Y, Z are uncorrelated random variables with zero means and S.D. 5, 12 and 9 respectively and if V = X + Y, W = Y+Z, find r_{VW} .**

10. **Three balls are drawn at random without replacement from a box containing 2 white, 3 red and 4 black balls. If X denotes the number of white balls drawn and *X* denotes the number of red balls drawn, find the joint probability distribution of (X, Y).**

11. **If X and Y are the RVs related to other two variables U and V such that *U = (X-a)/h* and $V = (y-b)/K.$ where *a,b,h* and *k* are constants and *h*, $k \neq 0,$ prove that $r_{XY} = r_{UV}$ (or) Prove that the correlation co-efficient is independent of the change of the origin and scale.**

12. **The joint probability distribution function of two random variables X and Y is**

$$f(x,y) = x+y \quad 0<x<1, \quad 0<y<1$$
$$= 0, \qquad \text{otherwise}$$

Find the correlation co-efficient between *x* and *y*

13. **Let X_1 and X_2 be two independent RVs with means 5 and 10 and S.D's 2 and 3 respectively. Obtain r_{UV} where $U = 3X_1 + 4X_2$ and $V = 3X_1 - X_2.$**

14. **For the following data taken from 10 observations, find out the regression equations of X on Y and Y on X: ΣX=250, ΣY = 300, ΣXY = 7900, ΣX^2 = 6500 and ΣY^2 = 10,000 Hence find *r*.**

15. **Find the co-efficient of correlation and obtain the lines of regression from the data given below;**

$x:$	62	64	65	69	70	71	72	74
$y:$	126	125	139	145	165	152	180	208

16. **For the following data, find the most likely price at Madras corresponding to the price 70 at Bombay and that at Bombay corresponding to the price 68 at Madras:**

	Madras	*Bombay*
Average price	65	67
S.D. of price	0.5	3.5

S.D of the difference between the prices at Madras and Bombay is 3.1.

17. **A two dimensional random variable (X,Y) have a bivariate distribution given by P(X=x, Y = y) = x^2 + y/32 for x = 0,1,2,3 and y = 0,1. Find the marginal distribution of X and Y.**

18. **If X and Y are independent variates uniformly distributed in (0,1) find the distribution of XY.**

19. Two random variables X and Y have the following joint pdf f(x,y) = 2 – x – y, $0 \le x \le 1$, $0 \le y \le 1$. Find the coefficient correlation.

20. If (X,Y) is a two – dimensional random variable uniformly distributed over the triangular region R bounded by y = 0, x = 3 and y = $\frac{4}{3}x$, find r_{XY}.

UNIT – III ESTIMATION THEORY

PART –A

1. **Define estimator, estimate and estimation.**
2. **Distinguish between point estimation and interval estimation.**
3. **Mention the properties of a good estimator.**
4. **Define confidence coefficient.**
5. **What is the level of significance in testing of hypothesis?**
6. **Define confidence limits for a parameter.**
7. **State the conditions under which a binomial distribution becomes a normal distribution.**
8. **Explain how do you calculate 95% confidence interval for the average of the population?**
9. **An automobile repair shop has taken a random sample of 40 services that the average service time on an automobile is 130 minutes with a standard deviation of 26 minutes. Compute the standard error of the mean.**
10. **What is a random number? How it is useful in sampling?**
11. **A population has the numbers: 12, 8, 10, 30, 12, 16, 40, 5, 16, 24, 22, 31, 30, 16, 12.**
12. **Draw a systematic sample of size 5. Find out its mean..**
13. **How large sample is useful in estimation and testing?**
14. **Define unbiasedness of a good estimator.**
15. **Let the lines of regression concerning two variables x and y be given by y = 32 – x and x = 13 – 0.25y. Obtain the values of the means.**
16. **What are the merits and demerits of the least square method.**
17. **Define maximum likelihood estimation.**
18. **Discuss the properties of maximum likelihood estimation**
19. **Find the MLE of f(x) = 1/θ, $0 \le x \le \theta$.**
20. **Find the MLE of $(\alpha+1)x^{\alpha}$, $0 < x < 1$.**

PART –B

1. **Fit a straight line y = ax +c to the following data.**

X	**1**	**3**	**5**	**7**	**9**	**11**	**13**	**15**	**17**
y	**10**	**15**	**20**	**27**	**31**	**35**	**30**	**35**	**40**

2. Find the regression line of Y on X for the data

x	**1**	**4**	**2**	**3**	**5**
y	**3**	**1**	**2**	**5**	**4**

3. In random sampling from normal population $N(\mu, \sigma^2)$, find the maximum likelihood estimator for μ when σ^2 is unknown.

4. x1, x2, x3,……xn are random observations on a Bernoulli variable X , taking the value 1 with probability θ and value 0 with probability $1-\theta$. Show that T(T-1)/n(n-1) is an unbiased estimate of θ^2 where T = sigma i=0 to n xi.

5. Let x1,x2…..xn denote a random sample from the distribution with pdf

$$f(x,\theta) = \theta x^{\theta-1}, \quad 0<x<1, \quad \theta>0$$

$$0 \qquad \text{Elsewhere}$$

prove that the product u1(x1,x2,….xn) = x1,x2…xn is a sufficient estimator for θ.

6. Let x1,x2,….xn be a random sample from uniform population on $[0,\theta]$. Find a sufficient estimator for θ.

7. Show that for a rectangular population $f(x,\theta) = 1/\theta, \quad 0 < x < \propto$

$$0 \quad \text{elsewhere}$$

Find the maximum likelihood estimator for θ.

8.For a random sampling from a normal population find the maximum likelihood estimators for

i) The population mean, when the population variance is known.

ii) The population variance, when the population mean is known.

iii)The simultaneous estimation of both the population mean and variance.

9.Obtain the lines of regression

X	**50**	**55**	**50**	**60**	**65**	**65**	**65**	**60**	**60**
Y	**11**	**14**	**13**	**16**	**16**	**15**	**15**	**14**	**13**

10. The price of a commodity during 93-98 are given below. Fit a parabola $y = a + bx + cx^2$ to these data. Calculate the trend values, estimate the period of the commodity for the year 1999.

x	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
y	**100**	**107**	**128**	**140**	**181**	**192**

UNIT – IV TESTING OF HYPOTHESIS

1. **What is the essential difference between confidence limits and tolerance limits?**
2. **Define Null hypothesis and Alternative hypothesis.**
3. **Define level of significance**
4. **Define Type-I error and Type-II error?**
5. **Define student's t-test for difference of means of two samples.**
6. **Write down the formula of test statistic't' to test the significance of difference betweenthe means.**
7. **Write the application of t-test?**
8. **What is the assumption of t-test?**
9. **State the important properties of 't' distribution.**
10. **Define chi square test of goodness of fit.**
11. **Define errors in sampling and critical region.**
12. **Write the application of 'F' test**
13. **Define a 'F' variate**
14. **A random sample of 25 cups from a certain coffee dispensing machine yields a mean** x **=6.9occurs per cup. Use _=0.05 level of significance to test, on the average, the machine dispenseμ=7.0 ounces against the null hypothesis that, on the average, the machine dispenses μ<7.0ounces. Assume that the distribution of** ounces per cup is normal, and that the variance **is theknown quantity** σ^2**=0.01 ounces.**
15. **In a large city A, 20 percent of a random sample of 900 school boys had a slight physicaldefect. In another large city B, 18.5 percent of a random sample of 1600 school boys had somedefect. Is the difference between the proportions significant?**
16. **A sample of size 13 gave an estimated population variance of 3.0 while another sample of size15 gave an estimate of 2.5. Could both samples be from populations with the same variance?**
17. **Give the main use chisquare. test.**
18. **What are the properties of "F" test.**
19. **Write the condition for the application of . test.**
20. **For a 2 x 2 contingency table**

a	b
b	d

write down the corresponding 2 value

PART – B(16 Marks)

1. (a) **A sample of 900 members has a mean 3.4 c.m and standard deviation 2.61 c.m. Is thesample from a large population of mean 3.25 c.ms and standard deviation of2.61c.ms?(Test at 5% L.O.S)**

   (b) Before an increase in excise duty on tea, 800 persons out of a sample of 1000 personswere found to be tea drinkers. After an increase in duty, 800 people were tea drinkers in a sample of 1200 people. Using standard error of proportion, State whether there is asignificant decrease in the consumption of tea after the increase in excise duty.

2. **(a) A manufacturer claimed that at least 95% of the equipment which he supplied to a factoryconformed to specifications. An examination of a sample of 200 pieces of equipmentrevealed that 18 were faulty. Test his claim at 5% level of significance.**

   (b) 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?

3. **(a) In a big city 325 men out of 600 men were found to be smokers. Does this informationsupport the conclusion that the majority of men in this city are smokers?**

   (b) Examine whether the difference in the variability in yields is significant at 5% L.O.S, for thefollowing.

	Set of 40 Plots	**Set of 60 Plots**
Mean yield per Plot	**1258**	**1243**
S.D. per Plot	34	28

4. (a) **The means of 2 large samples 1000 and 2000 members are 67.5 inches and 68.0 inchesrespectively. Can the samples be regarded as drawn from the same population ofstandard deviation 2.5 inches?**

(b) Two independent samples of sizes 8 and 7 contained the following values.
Sample I : 19 17 15 21 16 18 16 14
Sample II : 15 14 15 19 15 18 16
Test if the two populations have the same mean.

5. (a) Samples of two types of electric bulbs were tested for length of life and following data wereobtained.

	Type I	Type II
Sample Size	8	7
Sample Mean	1234hrs	1036hrs
Sample S.D	36hrs	40hrs

Is the difference in the means sufficient to warrant that type I is superior to type IIregarding the length of life?

(b) Two independent samples of 8 and 7 items respectively had the following Values of the variable (weight in kgs.)
Sample I : 9 11 13 11 15 9 12 14
Sample II: 10 12 10 14 9 8 10
Use 0.05 level of significance to test whether it is reasonable to assume that thevariances of the two population's sample are equal.

6. (a) A group of 10 rats fed on diet A and another group of 8 rats fed on diet B, Recordedthe following increase the following increase in weight.(gms)
Diet A : 5 6 8 1 12 4 3 9 6 10
Diet B : 2 3 6 8 10 1 2 8 - -
Does it show superiority of diet A over diet B ? (Use F-test)

(b) The marks obtained by a group of 9 regular course students and another group of 11 part timecourse students in a test are given below :
Regular : 56 62 63 54 60 51 67 69 58
Part-time: 62 70 71 62 60 56 75 64 72 68 66
Examine whether the marks obtained by regular students and part-time students differsignificantly at 5% and 1% levels of significance.

7. (a) Two independent samples of sizes 8 and 7 contained the following values.
Sample I : 19 17 15 21 16 18 16 14
Sample II : 15 14 15 19 15 18 16
Test if the two populations have the same variance.

(b) The average income of a person was Rs. 210 with S.D of Rs. 10 in a sample 100 peopleof a city. For another sample of 150 persons the average income

was Rs. 220 with S.Dof Rs. 12. Test whether there is any significant difference between the average incomeof the localities?

8. (a) Two random samples gave the following results:

Sample	Size	Sample mean	Sum of squares of deviation from the mean
1	10	15	90
2	12	14	108

Test whether the samples have come from the same normal population.

(b) Records taken of the number of male and female births in 800 families having four childrenare as follows :

Number of male births :	**0**	**1**	**2**	**3**	**4**
Number of female births :	**4**	**3**	**2**	**1**	**0**
Number of Families :	**32**	**178**	**290**	**236**	**64**

Test whether the data are consistent with the hypothesis that the binomial law holds and that the chance of a male birth is equal to that of female birth, namely p = ½ = q.

9. (a) Given the following table for hair colour and eye colour, find the value of Chi-square.Is there good association between hair colour and eye colour?

		Hair colour			
		Fair	Brown	Black	Total
Eye colour	Blue	15	5	20	40
	Grey	20	10	20	50
	Brown	25	15	20	60
	Total	60	30	60	150

(b) Out of 800 graduates in a town 800 are females, out of 1600 graduate employees 120 arefemales. Use chi squarȩ to determine if any distinction is made in appointment on the basis ofsex. Value of chi squarȩ at 5% level for 1 d.f is 3.84

10. (a) An automobile company gives you the following information about age groups

and the liking for particular model of car which it plans to introduce. On the basis of this data can it be concluded that the model aooeal is independent of the age group.

Persons who Below	20	20-39	40-59	60 and above
Liked the car	140	80	40	20
Disliked the car	60	50	30	80

(b) The following data gives the number of aircraft accidents that occurred during the variousdays of a week. Find whether the accidents are uniformly distributed over the week.

Days	Sun	Mon	Tues	Wed	Thu	Fri	Sat
No. of accidents	14	16	08	12	11	9	14

UNIT – V MULTIVARIATE ANALYSIS

PART A

1.Define random vectors and random matrices.

2. Define covariance matrix

3. State the properties of multivariate normal density.

4. Define Principal component analysis.

5. Define total population variance.

6. State the general objectives of principal components analysis.

7. State any two properties of multi variate normal distribution.

8. If X = $\begin{pmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \end{pmatrix}$ **Find** $\bar{X}$

9. If $\Sigma = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{pmatrix}$ Find the standard deviation matrix $V^{1/2}$

10. State the properties of covariance matrix

PART B

1. .Explain partitioning the covariance matrix.

2. Explain the mean vector amd covariance matrix for linear combination of random variables

3. Discuss Bivariate normal density.

4. Prove that the correlation coefficient between the components are the eigen values – eigen vector pairs for sigma.

5.. Consider the random vector $X' = \{X_1, X_2\}$. The discrete random variable X_1 have the following probability function.

X_1 : -1 0 1

$P_1(x_1)$: 0.3 0.3 0.4 and X_2 have the probability function

X_2 : 0 1

$P_2(x_2)$: 0.8 0.2 find the covariance matrix for the two random variables x_1 and x_2 when their joint pdf $p_{12}(x_1,x_2)$ is given by,

X1 / X2	0	1
-1	0.24	0.06
0	0.16	0.14
1	0.40	0

6. For the covariance matrix $\Sigma = \begin{pmatrix} 1 & 4 \\ 4 & 100 \end{pmatrix}$ the derived correlation matrix $P = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix}$, show that the principal components obtained from covariance and correlation matrices are different.

7. Let the random variables X_1,X_2,X_3 have the covariance matrix $\Sigma = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ determine the principal components Y_1,Y_2,Y_3

8. Let X_{3x1} be $N_3(\mu,\Sigma)$ with $\Sigma = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ Are X_1 and X_2 independent. What about (X_1,X_2) and X_3 .

9. Discussion of population principal components.

10 . Explain the various results of covariance matrix.

***** ALL THE BEST *****