VTM NSS COLLEGE DHANUVACHAPURAM

DEPARTMENT OF MATHEMATICS

QUESTION BANK

First semester complimentary Statistics for B.Sc Mathematics

ST1131.1:DESCRIPTIVE STATISTICS AND BIVARIATE ANALYSIS

Section B

1. Distinguish between census and sampling.
2. Define systematic sampling.
3. What is classification and tabulation of data.
4. Show that A.M≥G.M≥H.M for any set of data.
5. In a moderately asymmetrical distribution median is 41.6,mode is 48.4.Find mean.
6. Show that standard deviation is not affected by change of origin.
7. If the coefficient of variation of a distribution is 50 and its variance is 400.What will be the value of arithmetic mean.
8. Write the normal equation required for the fitting of the straight line y=ax+b.
9. Write the relationship between two variables X and Y is r=0.60.If the means and standard deviation of X and Y are 10,20,1.50 and 2.00 respectively,find the regression equation of Y on X.
10. Distinguish between positive and negative correlation.
11. What precautions are to be taken before we use a secondary data?
12. Define weighted arithmetic mean.
13. Show that the sum of deviations of the observations from the mean is zero.
14. Define quartiles. Define its importance.
15. Calculate the geometric mean of 2,4,8.
16. How do you compute the mean deviation about the mean in a discrete data.
17. What is the implication of positive and negative skewness?
18. For a moderately skewed distribution what is the relation between mean, median and mode?
19. How do you comment on the type of kurtosis with the help of the moment coefficient β_2
20. To fit a parabola of the form $y=ax^2+bx+c$,from a given data by the method of ordinary least squares, write the normal equations that are needed to estimate the parameters.
21. What are the underlying assumptions of Karl-persons' coefficient of correlation?
22. When do you go for computing rank correlation coefficient?
23. Why there are two regression lines? when do they coincide?
24. The class marks in a frequency table are given to be 5,10,15,20,25,30,35,40,45,50.Write down the classes. what are the considerations one has to bear in mind while forming a frequency distribution.
25. Explain any one method for collecting a primary data.
26. Distinguish between probability sampling and non-probability sampling.
27. What is SRSWOR?
28. What is sampling frame.
29. Mention any two measures of kurtosis.
30. Obtain the mean of first n natural numbers.
31. Give a situation where median is the most suitable average.
32. State some demerits of arithmetic mean

33. Calculate the standard deviation of 7,3,2,4,9,12,14.
34. What are the uses of scatter diagram?
35. Write the normal equations for fitting $y=ae^{bx.}$
36. Describe spearman's rank correlation coefficient.
37. What is meant by perfect correlation?
38. Write down the formula for both the regression equations.
39. What do you mean by raw and central moments?
40. What is the principle of least squares?

Section C

1. Distinguish between sampling and non-sampling errors.
2. Explain simple random sampling with replacement and without replacement.
3. Calculate the geometric mean for the following data.

Class:	1-3	4-6	7-9	10-12
Frequency:	8	16	15	3

4. Calculate the mean deviation about mean of 8,24,12,16,10,20.
5. The mean marks of 80 students of a class are 65.The mean of boys are 70 and that of girls is 62.Find the number of girls in the class.
6. Find the first,second and third moments about the origin for the set of numbers 1,3,5,7.
7. Explain the least square method of fitting of a parabola.
8. Calculate the rank correlation coefficient from the following data specifying the ranks of 7 students in two subjects.

   Rank in 1^{st} subject : 1 2 3 4 5 6 7

   Rank in 2^{nd} subject:4 3 1 2 6 5 7
9. Show that correlation coefficient is independent of change of origin and scale.
10. What is a simple random sample?Mention the various methods of drawing a random sample.
11. Briefly explain the construction of a histogram with example.
12. A sample of 35 values has mean 80 and standard deviation 4.A second sample of 65 values has mean 70 and standard deviation 5. Find the standard deviation of the combined sample of 100 values.
13. The mean of 5 items of an observation is 4 and the variance is 5.2.If three of the five items are 1,2,and 6.Find the other two.
14. What is kurtosis?How does it differ from skewness?
15. Calculate:

    a)The regression equation of X on Y and Y on X from the following data

    b)Estimate X when Y=20.

X :	10	12	13	17	18
Y :	5	6	7	9	13

16. Explain the fitting of a curve of the form y=a+bx.
17. Discuss the advantages of sampling over census.
18. Explain the concept of skewness and kurtosis.
19. Distinguish between simple random sampling and systematic sampling schemes.
20. Distinguish between nominal and ordinal scale of measurements.
21. Show that correlation coefficient is the geometric mean between the regression coefficients.
22. Explain the fitting of a curve $y=ax^2+bx+c$
23. Derive the angle between two regression lines.
24. Fit an equation of the form $y=ab^x$ to the following data.

X :	2	3	4	5	6

Y : 144 172.8 207.4 248.8 298.6

25. The two regression lines are y=ax+b and x=cy+d and the two variables have the same means.Then show that d(1-a)=b(1-c).
26. Form the regression line of Y on X for the given data
    X : 36 23 27 28 28 29 30 31 33 35
    Y : 29 18 20 22 27 21 29 27 29 28
27. Show that the correlation coefficient is invariant under linear transformation.
28. What are the steps in carrying out a statistical survey?
29. What are the advantages of stratification in a data?
30. The average of 100 workers was found to be Rs.600 .Later on it was discovered that the wages of two workers were misread as 350 and 450 instead of 300 and 500.Compute the correct mean of the data.
31. Write a comparison between absolute and relative measures of dispersion.What are the important absolute measures of dispersion?
32. Explain about cumulative frequency tables with examples.
33. What is mode.write down the procedure for finding mode.
34. What is the effect of the change in origin and scale on mean and standard deviation?
35. Find the variance of the first n natural numbers.
36. Derive the limits for correlation coefficient.
37. Explain the advantages of diagrammatic representation of data.
38. Construct a pie diagram for the following data giving the statistics of the IQ status of a group of students

Type of IQ:	Low IQ	Av.IQ	Good IQ	Exc IQ
No.of students:	20	15	10	5

39. The following table gives that distribution of marks of 100 students.If the mode of the distribution is 46.7,assuming that 40-50 as the model class,find the missing frequencies

Marks:	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No.of Students:	5	8	7	-	28	20	-	10

40. Define quartiles. What is their importance.?

Section D

1. Write briefly about raw moments and central moments.Derive the relation between them.

2. The following table gives the egg production during a year at a poultry farm.

No.of eggs:	*0-30*	*30-60*	*60-90*	*90-120*	*120-150*	*150-180*	*180-210*	*210-240*
No.of hens:	*3*	*4*	*12*	*33*	*69*	*92*	*50*	*25*

No.of eggs:	*240-270*	*270-300*
No.of hens:	*11*	*1*

   Represent the above frequency distribution graphically(histogram).
   Also draw the ogives and estimate median value.

3. Compute the pearson's moment measure of skewness and kurtosis for the following data and comment on the nature of the distribution based on your findings from the following data.

Wages : *10-12 12-14 14-16 16-18 18-20 20-22 22-24*

No.of workers: *1 3 7 20 12 4 3*

4. a)what is spearman's rank correlation coefficient
   b)The scores for nine students in physics and math are as follows.Compute the spearman's rank correlation coefficient and comment on your finding.

   Physics: *35 23 47 17 10 43 9 6 28*

   Mathematics: *30 33 45 23 8 49 12 4 31*

5. a) Discuss the various types of correlation
   b)For the following data fit a straight line of the form y=a+bx
   Also estimate y when x=20

   X : 0 1 2 3 4
   Y : 0 1.8 3.3 4.5 6.3

6. Construct a histogram and frequency polygon for the following data

Class interval	5-9	10-14	15-19	20-24	25-29	30-34	35-39
Frequency	8	15	18	30	16	12	6

   Draw an ogive and estimate the median.

7. Calculate
   (i) Quartile deviation and
   (ii) Mean dxedviation from mean,for the following data

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No of students	6	5	8	15	7	6	3

8. Draw the frequency curve to the following data

X :	0-20	20-40	40-60	60-80	80-100
F :	9	15	23	10	8

   Draw ogives and then find quartile deviation.

9. Fit a curve of the form $y+a+bx+cx^2$ to the following data

   x : 10 15 20 25 30 35 40
   f : 11 13 16 20 27 34 41

10. (a)Derive the standard error in the estimate of y in regression of y on x.

    (b)Show that correlation coefficient , $(r)=\frac{\sigma_x^2+\sigma_y^2+\sigma_{x-y}^2}{2\sigma_x\sigma_y}$ where σ_x^2, σ_y^2 and σ_{x-y}^2 are the variances x,y and x-y respectively.

11. Calculate the rank correlation coefficient for the following data on heights of fathers and sons.

    Ht of father : 65 63 67 64 68 62 70 66 68 67
    Ht of son : 68 66 68 65 69 66 68 65 71 67

12. (a)Given that the two regression equations are 8x-10y+66=0 and 40x-18y-214=0.Obtain regression coefficients and the correlation coefficient.Find the means of X and Y.
    If the standard deviation of x is 4, find the standard deviation of y.

    (b) Derive the regression line of x on y.

13. fit a parabola to the following data.

    X : 1 2 3 4 5 6 7 8 9

    Y : 2 6 7 8 10 11 11 10 9

Estimate y when x=4.5.

14. Ten competitors in a beauty contest are ranked by three judges in the following order.

I judge :	1	5	4	8	9	6	10	7	3	2
II judge :	4	8	7	6	5	9	10	3	2	1
III judge :	6	7	8	1	5	10	9	2	3	4

Use rank correlation to discuss which pair of judges have the nearest approach to common tastes in beauty.

15. Calculate the arithmetic mean and the median of the frequency distribution given below.Hence calculate the mode using the empirical relation

Classes:	130-134	135-139	140-144	145-149	150-154	155-159	160-164
F:	5	15	28	24	17	10	1

16. (a)Explain the advantages of diagrammatic representation of data.

(b)Construct a pie diagram for the following data giving the statistics of the IQ status of a group of students

Type of IQ:	Low IQ	Av.IQ	Good IQ	Exc IQ
No.of students:	20	15	10	5

17. (a) Show that correlation coefficient is the geometric mean between the regression coefficients

(b) Derive the angle between two regression lines.

18. Construct a histogram and frequency curve for the following data

x:	10-12	12-14	14-16	16-18	18-20	20-22	22-24
f:	1	3	7	20	12	4	3

19. (a)Explain the fitting of a curve $y=ax^2+bx+c$

(b) Fit an equation of the form $y=ab^x$ to the following data.

X :	2	3	4	5	6
Y :	144	146	200	250	300

20. Form the regression line Y on X and X on Y for the given data

X :	28	28	29	36	23	27	30	31	33	35
Y :	21	29	27	29	28	29	18	20	22	27