NORTH MAHARASHTRA UNIVERSITY JALGAON

QUESTION BANK

FOR

F.Y.B.Sc

STATISTICS

PAPER-I

DESCRIPTIVE STATISTICS

(With effect from June 2007)

F.Y.B.SC.

PAPER-I   DESCRIPTIVE STATISTICS

Chapter – 1   Introduction to basic concepts

A) Questions of 2 marks

1. Define statistics

2. What do you mean by population and statistical population?

3. What is a sample?

4. Draw a simple random sample with replacement of size n = 2, if the population units are 2, 3, 5.

5. Define sampling frame

6. In a population of size N = 6, the observations were 3, 4,7,9,10,11. Draw all possible SRSWOR of size 2.

7. If a population consists of 30 items then how many

   I) SRSWOR each of size 10   II) SRSWR each of size 10 can be
    selected

8. Explain with illustration the term

   i) Finite population   ii) infinite population

9. The word statistics has been derived from the Latin word ………. Or the German word …….

10. Two Indian statisticians who have made significant contribution in the development of statistics are ………………….and ……………

11. Comment on the statement

    "Statistics is the science of counting"

12. Indicate whether the following statement is True or False

    i)   Statistics is of no use to humanity

    II)   To a very striking degree, our culture has become a statistical culture

13. Give illustrations of each of the following sampling methods

    i) SRSWR   ii) SRSWOR

14. Give illustrations of each of the following methods

     i)   Stratified random sampling

     II)   Systematic sampling

B) Questions for 4 marks

1. Discuss the scope of statistical methods
2. Write a note on statistical organizations in India with special reference to NSSO.
3. Describe central Statistical organization. CSO.
4. Explain in brief Indian Statistical Institute ISI.
5. Describe the functions of Bureau of economics and statistics.
6. Explain the method of simple random sampling.
7. Describe the method of stratified random sampling
8. What do you mean by systematic sampling? Explain briefly.
9. State the advantages of sampling over census.
10. How does SRSWR differ from SRSWOR?
11. Make critical comparison between sampling and census
12. Distinguish between Random sampling and Non-random sampling
13. Explain the situation where sampling has larger scope as compared to census.
14. From the population containing 20 units obtain a systematic sample of size 5. Suppose 4th unit is the first unit selected at random in the systematic sample. Also obtain all possible systematic samples.
6,18,10,12,16,14,15,13,9,13,12,17,16,15,8,12.
15. State the disadvantage of sampling over census.
16. Explain the situations where sampling has larger scope as compared to census.

C) Questions for 6 marks

1. Describe the importance and limitation of statistics
2. Discuss the advantages and disadvantages of sampling over census
3. Explain what is random sample. Why random sample is referable?
4. Explain the various methods of achieving randomness.

5. What are the requirements of a good sample?

6. Explain the terms; population, sample and sampling unit.

Chapter- 2 Data And Measurement Scales

A) Questions for 2 marks

1. Define a variable.

2. What do you mean by an attribute?

3. Give an illustration of Nominal scale where it is used?

4. Define the term ordinal scale

5. Explain the term Nominal scale

6. Define the term Interval scale

7. Explain the term Ratio scale

8. Give Practical application of ordinal scale

9. Give an example of ratio scale where it is used?

10. Define discrete variable

11. Define continuous variable

12. Use appropriate scale in following

    i)   Classification of students in various divisions of same standard.

    ii)   Groups of students according to grades in examination such as fail, second class, first class, distinction.

13. What is raw data?

14. Fill in the blanks

    Data are classified into. …………and …………

15. Fill in the blanks

    Data originally collected for any investigation is called……………..and the …………….data should be used after careful scrutiny.

B) Questions for 4 marks

1. Write a note on various scales of measurement.

2. Describe the four different scales of measurements

3. Distinguish between 'nominal scale' and 'ordinal scale' with illustrations

4. Distinguish between 'interval scale' and 'ratio scale' giving suitable illustrations.

5. Distinguish between primary and secondary data with suitable examples

6. Write a note on sources of secondary data.

7. Describe the meaning of data and statistical data

8. How to collect raw data, explain with suitable illustrations.

C) Questions for 6 marks

1. Distinguish between Attributes and Variables

2. Explain with illustrations Qualitative and Quantitative data

3. Explain the different methods of collecting primary data

4. What are the various methods of collecting statistical data? Which of these are most reliable and why?

5. Describe the methods generally employed in the collection of statistical data, stating briefly their merits and demerits.

6. Distinguish between primary data and secondary data and discuss the various methods of collecting primary data

7. Distinguish between primary data and secondary data. What precautions should be taken in the use of secondary data?

Chapter- 3 Presentation of Data

A) Questions for 2 marks

1.Explain the need of classification

2. Define the term tabulation

3. Explain the purposes of population

4. State any two objectives of table

5. Define the term table with illustration

6. Define the term Class frequency with illustration

7. Define the term open end classes with illustration

8. Define the term class width with illustration

9. Define the term mid value with illustration

10. Define the term class boundaries with illustration

11. Define the term class limits with illustration

12. Define the term less than cumulative frequency with illustration

13. Define the term more than cumulative frequency with illustration

14. To represent the data related to a single variable, we used -

   a) Simple bar diagram   b) Multiple bar diagram

   c) Subdivided bar diagram   d) Pie diagram

15. To represent the data by a circle, a diagram we used –

   a) Simple bar diagram   b) Multiple bar diagram

   c) Subdivided bar diagram   d) Pie diagram

16. To draw histogram, height of rectangle for unequal class width is

   Proportional to

   a) Class frequency   b) class width

   c) Frequency density   d) mid values

17. What are the uses of histogram?

18. What is a frequency distribution of a variable?

19. Which diagram is appropriate for the following situation

   "Average yield per acre of wheat in India for the years from 1981 to

   1990."

20. Which diagram is appropriate for the following situation

   "Number of students enrolled in a college in three faculties of Arts,

   Science and Commerce for the years 1987-88, 1988-89 and 1989-90."

21. Which diagram is appropriate for the following situation?

   "Cost of construction of building under different heads of labour, bricks,

   steel, cements, timber and supervision."

B) Questions for 4 marks

1. Explain the different methods of classification

2. Distinguish between – Inclusive classes and Exclusive classes.

3. Prepare a blank table giving the following information about workers in a
   certain industry

   Sex   : Male, Female

   Age group: 20 – 30, 30 – 40, 40 and above

   Skill   : Skilled, Unskilled

4. What are the advantages of tabulation?

5. What are the requirements of a good table?

6. Present the following information in a tabular form by computing the figures which are not given.

   "Out of 800 employees appeared for a promotion test, 320 were married. Among 240 who were unsuccessful, 96 were married."

7. Complete the following table showing data related to examination result.

	F. Y.	S. Y.	T. Y.	Total
First class	148	82	---	---
Second class	192	95	38	---
Pass class	---	108	---	210
Fail	20	---	90	150
Total	---	325	275	1000

8. Explain the general guidelines of preparing a frequency distribution.

9. Sate the advantages of graphical representation of data.

10. Sate the advantages of diagrammatic representation of data.

11. Explain the construction of pie diagram with rough sketch.

12. Explain the construction of Simple bar diagram with rough sketch.

13. Explain the construction of Multiple bar diagram with rough sketch.

14. Explain the construction of Subdivided bar diagram with rough sketch.

15. Represent the following data by simple bar diagram.

Years	1981	1982	1983	1984	1985
Production (in million tones)	45	40	50	52	47

16. Represent the following data by multiple bar diagram.

Years	Exports (in cr.of Rs.)	Imports (in cr.of Rs.)
1993 – 94	320	250
1994 – 95	340	260
1995 – 96	340	240
1996 – 97	310	200

17. Represent the following data by subdivided bar diagram.

Quarters	Production		Total
	Finished Articles	Incomplete Articles	
I	300	200	500
II	350	200	550
III	400	150	550
IV	450	100	500

18. Draw a Pie diagram to represent the following data.

Source	Percentage Receipts
Individual income taxes	45
Corporation income taxes	15
Social insurance taxes & contributions	30
Excise taxes	05
Others	05

19. Exhibit the following data by using Stem & leaf chart.

    10.8   12.9   13.0   12.5   13.1   10.4   10.2   13.3   11.8   11.1

    11.3   11.6   11.2   11.1   11.3   11.6

20. Explain the construction of Ogives along with the rough sketch.

21. Explain the steps of procedure of classification of continuous variable.

C) Questions for 6 marks

1. Explain the different parts of a statistical table.
2. Explain the different types of table.
3. Present the following information in a tabular form by computing the figures which are not given.

   "In a certain interview there were 150 candidates of which 56 % were males. 36 candidates were successful in the interview. The proportion of males to females in the successful candidates is 5:4."

4. Explain the construction of following graphs along with the rough sketches.

   i) Frequency curve   ii) Frequency polygon

Chapter- 4 Measures Of Central Tendency

A) *Questions for 2 marks*

1. What do you mean by central tendency.

2. We can determine median graphically by

   a) Frequency curve   b) Histogram

   c) Less than C.F.curve   d) Frequency polygon

3. We can determine mode graphically by

   a) Frequency curve   b) Histogram

   c) Less than C.F.curve   d) Frequency polygon

4. We can determine partition values graphically by

   a) Frequency curve   b) Histogram

   c) Less than C.F.curve   d) Frequency polygon

5. Define weighted A. M.

6. Define weighted G. M.

7. Define weighted H. M.

8. Show that the algebraic sum of deviations of observations from their A. M. is zero.

9. Find A. M. given that $\sum(x - 10) = 230$ and $n = 50$

10. G. M. of 2, 8, 20, 0, 55 is

    a) 15.85   b) 0   c) 5.85   d) 10.85

11. The median of 35, 38, 40, 39, 35, 36, 37 is

    a) 38   b) 36   c) 37   d) 36.5

12. State the empirical relation between A. M., Median, and Mode.

13. Explain the term trimmed mean.

14. Mention the situations in which the G. M. useful as average

15. Mention the situations in which the H. M. useful as average

16. For a moderately asymmetric distribution mean = 25 and median = 23. Obtain the mode of the distribution.

17. State the various measures of central tendency.

18. If A. M. =   25   H. M. = 10   then G. M. =

19. If A. M. = 30   median = 26   then Q2 =

20. If $\sum Xi = 525$ A. M. =25   then n =

B) Questions for 4 marks

1. State requirements of a good measure of central tendency.
2. Discuss merits and demerits of A. M.
3. Discuss merits and demerits of Median.
4. Discuss merits and demerits of Mode.
5. Show that the sum of squares of deviations of all observations taken from A. M. is minimum.
6. Derive formula for combined A. M. of k groups given their sizes and A.M's.
7. Discuss the effect of change of origin and scale on A. M.
8. Find A.M. of first n natural numbers.
9. A variable takes values 1, 4, 9… n^2, find its A.M.
10. A variable takes values 1, 4, 9… n^2 with frequencies 1, 2, 3… n respectively. Find its A.M.
11. The mean monthly salary paid to 300 employees of a firm is Rs. 1470. There are 200 male employees & the remaining are females. If mean salary of males is Rs. 1505, obtain the mean salary of females.
12. A.M. of 50 items is 104. While checking, it was noticed that observation 98 was misread as 89. Find the correct mean.
13. If a & b are any two positive observations then prove that the corresponding A.M.$\geq$G.M. $\geq$H.M.
14. If a & b are any two positive observations then prove that
    G.M. = (A.M. x H.M.)½.
15. State the uses of Box plot.
16. A train travels the first half distance of its journey with a speed of 25 km./hr., the next one-fifth distance of a speed of 60 km./hr., the remaining one fourth distance at a speed of 40 km./hr. Find the average speed of the train during the journey.
17. A cost of machine is depreciated by 15%, 20%, & 30% in the first year, second year and third year respectively. Find the average percentage of depreciation.
18. Mean daily salary of 50 employees in a firm is Rs. 88.40 Frequency distribution of salaries of these employees in which some frequencies are missing, find it.

Salary	40 – 60	60 – 80	80 – 100	100 – 120	120 – 140
Frequency	6	---	17	---	5

19. A variable takes values 0, 1, 2… n with frequencies $^nC_0, {}^nC_1, {}^nC_2,$ ---- ^nC_n respectively. Find its A.M. of X.

20. Represent the following data using Box plot technique.

15, 11, 11, 28, 22, 14, 13, 18, 30, 12, 25

21. A.M. of the following frequency distribution is 5, find the value of x

Variable	2	4	6	8
Frequency	x – 1	x+1	x+1	2x - 5

22. State the utility of measures of central tendency.

23. A variable takes values 7, 10, 13 … 50 terms, find its A.M. & median.

C) *Questions for 6 marks*

1. Define A.M., Median, Mode and state the formula for each, in case of individual observations and frequency distribution.

2. If a,b, and c are any three positive observations then prove that the corresponding A.M.$\geq$G.M. $\geq$H.M.

3. A variable takes values a, ar, ar²…, ar^{n-1}, find it's A.M. and G.M.

4. A set of 10 values has A.M. 20. Find the A.M. if

   i) Each value is doubled & then increased by 2.

   ii) Each value is increased by 5 & then doubled.

5. Define Quartiles, Deciles, and Percentiles. Give the formula each for frequency distribution.

Chapter 5 - Measures Of Dispersion

A) *Questions for 2 marks*

1. The measurement bared on all values of the series is

   a) Range   b) Standard Deviation

   c) Quartile Deviation   d) All the above

2. In case of open class intervals, more suitable measurement of dispersion is

    a) Mean   b) Standard Deviation

    c) Quartile Deviation   d) None of these

3. The measurement least affected by extreme values of the series is

    a) Range   b) Standard Deviation

    c) Quartile Deviation   d) Mean Deviation

4. Interquartile range includes

    a) first 50% of the data   b) Last 50% of the data

    c) Middle 50% of the data   d) More of these.

5. Quartile deviation is

    a) $Q_3 - Q_1$ b) $Q_1 + Q_3$

    c) $\dfrac{Q_3 - Q_1}{2}$ d) $\dfrac{Q_1 + Q_3}{2}$

6. Mean deviation is minimum about

    a) Mean   b) Median

    c) Mode   d) First Quartile

7. The mean of 200 items is 48 & S.D. is 3. The sum of squares of there items is

    a) 46, 260   b) 4,62,600

    c) 46, 200   d) 1,54,200

8. If N=10, $\sum X = 60$, $\sum X^2 = 1000$, then S.D. is

    a) 8   b) 12

    c) 6   d) 100

9. If $\sigma = 16$, $\sum (x - \bar{x})^2 = 4096$ then the value of N is

    a) 16   b) 256

    c) 4   d) None of these

10. The formula for coefficient of variation is

    a) $\dfrac{S.D.}{Mean}$ b) $\dfrac{Mean - Mode}{S.D.}$

    c) $\dfrac{S.D.}{Mean - Mode}$ d) $\dfrac{Mean}{S.D.}$

11. If S.D. is 4, number of items is 10 & the sum of items is 160, the coefficient of variation is

    a) 16%   b) 25%

    c) 20%   d) 35%

12. The mean of a frequency distribution is 100 & coefficient of variation is 45% then standard deviation will be

    a) 45   b) 0.45

    c) 4.5   d) 450

13. The scale on X-axis in Lorenz curve is written as

    a) 100 to 0   b) 10 to 100

    c) 1000 to 100   d) None of these

14. Algebraic sum of deviation from arithmetic mean is

    a) Positive   b) Negative

    c) Zero   d) Different for each other

15. Sum of squares of deviations is minimum when taken from

    a) Mean   b) Median

    c) Mode   d) None of these

16. Sum of absolute deviations is minimum when measured from

    a) Mean   b) Median

    c) Mode   d) None of these

17. For a discrete frequency distribution,

    a) S.D. $\leq$ M.D. about mean   b) S.D.$\geq$ M.D. about mean

    c) S.D.= M.D. about mean   d) S.D. < M.D. about mean

18. The range of a given distribution is

    a) Greater than S.D.   b) Less than S.D.

    c) Equal to S.D.   d) None of these

19. Mean deviation from median is given by

    a) $\dfrac{\sum |x - md|}{n}$ b) $\dfrac{\sum |x - md|}{\sqrt{n}}$

    c) $\sqrt{\dfrac{\sum |x - md|^2}{n}}$ d) $\dfrac{\sum |x - md|^2}{n}$

20. For the five observations 7,7,7,7,7

    a) Mean=7 Variance=7   b) Mean=0 Variance=0

    c) Mean=7 Variance=0   d) Mean=0 Variance=49

B) Questions for 4 marks

1. A variable takes values 2,3,4,…10. Find the mean deviation about median.
2. A variable takes values 1,2,…n. Find the mean and variance.
3. Two samples of sizes 40 & 50 have the same mean but different standard deviations 19 & 8 respectively. Find the standard deviation of the combined group.
4. If n=100, $\sum x = -20$, $\sum x^2 = 220$. Find S.D. & C.V.
5. Coefficients of variation of the two series are 60% & 80%. Their standard deviations are 20 & 16. What are their arithmetic means?
6. The mean & S.D. of 100 observations are 50 & 10 respectively. Find the new mean & S.D. if

   i) if 2 is added to each observation

   ii) if 3 is subtracted from each observation

   iii) if each observation is multiplied by 5

   iv) if 2 is subtracted from each observation & then it is divided by 5
7. For a group of 30 male workers, the mean & S.D. of weekly overtime work in hours are 10 & 4 respectively, for 20 female workers the mean & S.D. are 5 & 3 respectively.

   i) Calculate the mean for the two groups taken together.

   ii) Is the overtime work more variable for the male group than for the female group? Explain.

C) Questions for 6 marks

1. What is dispersion? Give objectives of dispersion.
2. Define mean deviation. Prove that mean deviation is least when it is taken about median.
3. Define mean square deviation. Prove that mean square deviation about mean is the least among all mean squared deviations.
4. Define variance. Prove that variance is not affected by change of origin.
5. Define variance. If Y=KX prove that $V(Y) = K^2 V(X)$.
6. Define standard deviation. Why standard deviation is regarded as superior to other measures of dispersion.

7. What is coefficient of variation? Which purpose does it serve?

8. Explain relative & absolute measures of dispersion.

9. Distinguish between absolute & relative measures of dispersion.

10. Define Lorenz curve. What are its merits and demerits?

11. Explain the method of construction of Lorenz curve.

12. Let $x_1, x_2, \ldots x_n$ be n observations and let $\bar{x}$ be their arithmetic mean. Prove

   i) $\sum \left(x_i - \bar{x}\right) = 0$

   ii) If $u_i = \dfrac{x_i - a}{h}$ where a & $h \neq 0$ are constants, then S.D.(X) = h .S.D(U).

   iii) $\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2 \leq \sum_{i=1}^{n} (x_i - a)^2$ Where 'a' is a constant.

13. Derive the formula for combine group variance.

14. Two groups with n_1 & n_2 items have the same mean but different variances

   S_1^2 & S_2^2. Show that the variance S^2 of the combined group is given by

   $$S^2 = \frac{1}{n_1 + n_2} \left[n_1 S_1^2 + n_2 S_2^2\right]$$

15. Prove that for any discrete distribution the mean deviation about mean is less than the S.D.

16. If the mean & standard deviation of a variable X are m & σ respectively.

   Obtain the mean & standard deviation of $\dfrac{ax+b}{c}$ Where a, b, c are constants.

17. If X_i, i=1, 2… n are observations on X. Show that $\sum X_i^2 \geq \dfrac{\left(\sum X_i\right)^2}{n}$.

Chapter - 6   Moments Skewness and Kurtosis

A) *Questions for 2 marks*

1. If the mean, mode and s.d. of a frequency distribution are 41, 45 and 8 respectively, then the Pearson's coefficient of skewness is :

   a) $-\dfrac{1}{2}$ b) $\dfrac{1}{2}$ c) (- 1)   d) + 1.

2. What is symmetric distribution? μ_{2n+1} =   -------

a) 0   b) 1   c) – 1   d) 0.5

3. What is the relation between β_1 and β_2?

4. If the kurtosis of a distribution is 3, it is called ----- distribution and if it is 0, it is called ---- distribution.

5. For a symmetric distribution

    a) $\mu_2 = 0$ b) $\mu_2 > 0$ c) $\mu_3 > 0$ d) $\mu_3 = 0$.

B) Questions for 4 marks

1. In a certain distribution upper quartile exceeds the median by 10 units whereas the median exceeds the lower quartile by 7 units. Compute the coefficient of skewness.

2. For a frequency, Bowley's coefficient of skewness is 0.6. The sum of first and third quartiles is 100 and the median is 38, Find the two quartiles.

3. Find the c.v. of a frequency distribution given that its mean is 120, mode is 123 and Karl Pearson's coefficient of skewness is -0.3.

4. The first two moments of a distribution about the value 4 are 3 and 34. Find the mean and variance.

5. What is the effect of change of origin and scale on moments

6. If $u = kx$, show that μ_r of $u = k^r \mu_r$ of x.

7. Show that the measures of skewness are independent of both change of origin and scale.

8. For a symmetric distribution, with usual notations prove that,

$$\mu_3' = 3\mu_2' + \mu_1'^2$$

9. For a frequency distribution, Bowley's coefficient of skewness is 0.6. The sum of the first and third quartiles is 100 and the median is 38. Find the two quartiles.

10. Define 'kurtosis'. What are the types of kurtosis?

11. Show that $\beta_2 \geq 1$, with usual notations.

12. The first four moments about 4 of a certain distribution are 1.5, 17, -30 and 308. Find kurtosis and interpret.

13. Given that $\beta_2 = 2.6$, $\beta_1 = 0.19$, $\mu_2 = 1.2$ Find μ_3 and μ_4.

14. Explain the term 'skewness' using suitable diagrams.

15. Explain the different measures of skewness.

16. For two distributions A and B following summary statistics are available.

	A	B
Median	20	24
Q_1	13	14
Q_3	30	31

Compare the skewness of two distributions using appropriate measures of skewness.

17. Given that $\bar{x} = 1$, $\mu_2 = 3$, $\mu_3 = 0$ and $\mu_4 = 27$ find the first four raw moments.

18. Express first four central moments in terms of raw moments.

19. Show that for any frequency distribution

    i) Kurtosis is grater than unity.

    ii) Karl-Pearson's coefficient of skewness lies between -3 and +3.

20. Find the third central moment of the following observations

    The first four moments of a distribution about the value 4 of the variable are

    -1.5, 17, -30, and 108. Find the moments about the mean.

C) Questions for 6 marks

1. Define raw and central moments of (i) frequency distribution .

2. The first four moments of a distribution about the value '5' are 2, 20, 40, and 200 respectively. Find the first four central moments.

3. A distribution has mean 30, coefficient of variation 20% and coefficient of skewness is 0.3. Find its mode.

4. In a certain frequency distribution the sum of upper and lower quartiles is 45 and the difference between them is 15 . If the median is 20, find the coefficient of skewness

5. Define Karl-Pearson's coefficients γ_1 and γ_2 and discuss their utility in Statistics.

6. Given that n = 100, $\Sigma x^2 = 140$ $\Sigma x^3 = -40$, $\Sigma x^4 = 560$. Find β_1, β_2 and comment on the nature of skewness and kurtosis of the distribution.

7. Given that A.M. = 160 , mode = 157, 6 = 50 , find (i) Karl Pearson's coefficient of skewness (ii) median (iii) coefficient of variation.

8. Define raw and central moments of series of individual observation.

Chapter 7 - Correlation

A) *Questions for 2 marks*

1. Define the term correlation.
2. What is bivariate data?
3. State the merits of scatter diagram.
4. Discuss demerits of scatter diagram.
5. Which are measures of correlation?
6. Define covariance.
7. Prove that $Cov(X,X)=Var(X)$.
8. Given X = 53, Y = 28, byx = -1.5 & bxy = - 0.2. Find r
9. Explain the terms-   a) rank   b) tie of rank
10. State the properties of covariance.
11. When two variables said to be correlated? Give two examples.
12. What do you mean by positive correlation? Give two examples.
13. What do you mean by negative correlation? Give two examples.
14. If r = 0. Comment on it.
15. If r = 0 → covariance = 0 comment.
16. If covariance = 0 → r = 0 comment.
17. If the correlation co-efficient between two variables x & y is positive then comment on the following –
    a) The correlation coefficient between -x & -y is positive.
     b) The correlation coefficient between -x & y or x & -y is positive.

B) *Questions for 4 marks*

1. Comment on "If the correlation coefficient between two variables is zero then the variables are independent.
2. Comment on "'r 'measures every type of relationship between the two variables". Justify your answer.
3. Comment on "When the coefficient of correlation r between the variables is zero. The rank correlation is also zero & the variables are independent.

4. Karl Pearson's coefficient of correlation of two variables x & y is 0.8. Their covariance is 40. If the variance of x series is 16, find the standard deviation of y series.

C) Questions for 6 marks

1. Define Karl Pearson's coefficient of correlation & show that it lies between -1 & 1.
2. What is rank correlation? Derive formula for the Spearman's rank correlation coefficient.
3. Define product moment correlation coefficient & give the interpretation of different values of it.
4. Show that Spearman's rank correlation coefficient lies between -1 & 1.
5. Explain the scatter diagram. Describe the use of scatter diagram & limitations of scatter diagram.
6. The coefficient of rank correlation between marks in statistics & marks in mathematics obtained by a certain group of students is 0.8. If the sum of the squares of the difference in ranks is given to be 33, find the number of students in the group.
7. The coefficient of rank correlation of the marks obtained by 10 students in mathematics & statistics was found to be 0.5. It was latter discovered that the difference in ranks in two subjects obtained by one of the students was wrongly taken as 3 instead of 7. Find the correct coefficient of rank correlation.
8. From the following data, compute the correlation between X & Y.

	X Series	Y Series
No. of items	15	15
Arithmetic Mean	25	18
Sum of squares of deviations from mean	136	138

Summation of product of deviations of X & Y series from the respective arithmetic means=122.

9. In two sets of variables X & Y with 50 observations each, the following data were observed:

$\overline{X}$ = 10, σ_x =3, $\overline{Y}$ =6, σ_y =2 & r(X,Y,)=0.3

But on subsequent verification it was found that one value of X (=10) & one value of Y (=6) were inaccurate & hence weeded out. With remaining 49 pairs of values, how is the original value of r affected?

Chapter- 8   Regression

A) *Questions for 2 marks*

1. Given b_{yx} = - 1.4 & b_{xy} = - 0.5, calculate r_{xy}.
2. Comment on the following. For a bivariate distribution,b_{yx} = 4.2 & b_{xy} = 0.5.
3. For a bi-variate distribution b_{yx} = 2.8 and b_{xy} = - 0.3 Comment.
4. With b_{xy} = 0.5, r = 0.8 and variance of y = 16, find σ_x .
5. A student obtained the two regression lines as 2x – 5y – 7= 0 & 3x + 2y – 8= 0
    Do you agree with him ?
6. Comment on the following :-
    The correlation coefficient between x & y is 0.90 & one regression
    coefficient b_{xy} is – 1.
7. If b_{yx} = 0.8 & r_{xy} = 0.69, what would be the value of b_{xy} ?
8. "The regression equations of y on x & x on y are irreversible" Explain.
9. A correlation coefficient r = 0.8 indicate a relationship twice as close as
    r = 0. 4" comment.
10. What is the standard error of estimating y from x if r = 1?
11. What is the standard error of estimating y from x if r = 0?
12. When are regression lines i) perpendicular to each other &   ii) coincide ?
13. What is the angle between two regression lines if
    i)   r = + 1
    ii)   r = - 1
    iii)   r = 0.
14. comment on the following "The regression coefficients are symmetric
    functions of x and y".
15. Show that the A.M. of two regression coefficients is garter than the
    correlation coefficient.

B) Questions for 4 marks

1. What do you mean by regression? Why are there two regression lives in case of a bio-variate series ?
2. When are the regression lines (i) perpendicular to each other & (ii) coincide?
3. Define regression coefficients. What information do they supply ?
4. Explain the least square principle of obtaining regression lines.
5. Distinguish between regression coefficients & correlation coefficient.
6. Explain the concept of 'explained variation' and ' unexplained variation' of dependent variable.
7. Define coefficient of determination & state its utility.
8. Show that regression lines intersect at ($\bar{x}$, $\bar{y}$) .
9. Explain the procedure of fitting curve $y = a+bx+cx^2$.
10. Explain the procedure of fitting curve $y = ab^x$.
11. Find out the regression coefficients of Y o X and X on Y on the basis of following data

    $\bar{Z} X = 50$, $\bar{X} = 5$, $\bar{Z} Y = 60$, $\bar{Y} = 6$, $\bar{Z} x y = 350$,

    Variance of x = 4, variance of y = 9.
12. The lines of regression of a bivariate population :-

    8x – 10y + 66 = 0

    40x – 18y = 214

    Find i) the mean values of x & y ii) correlation coefficient between x & y
13. A student obtained the two regression lines as

          2x – 5y – 7 = 0

    &   3x + 2y – 8 = 0

    Do you agree with him?

C) Questions for 6 marks

1. Derive an expression for the acute angle between the two regression lines and discus the cases i) r = 0 ii) r = 1 .
2. With usual notation, prove that a) byx = r^2 b) byx & bxy can't exceed unity simultaneously c) (byx + bxy) $\geq r$, provide r > 0 .
3. Derive standard error of regression estimate of y on x .
4. Derive the expression for regression line of y on x .

5. The correlation coefficient between x & y is r = 0.60 If 6x = 1.50

   6y = 2.00, $\bar{x}$ = 10   & $\bar{y}$ = 20, find the equations of regression

   lines i) y on x & ii) x on y .

6. Given x = 4y + 5 & y = kx + 4   are the lines of regression of x on y and y

   on x respectively. If k is positive, prove that it can't exceed ¼ . two

   variables & coefficient of correlation between them.

7. The lines of regression of y on x and x on y are y = 0.3x + 10 .0 and

   x = 1.2y + 0.8 respectively. Determine the means of x & y the ratio of the

   standard deviation of x & y, the correlation coefficient between x & y.

Chapter- 9 Attribute

A) Questions for 2 marks

1. Define the term   dichotomy & manifold classification

2. Comment on the association between the attributes A & B in each of the

   Following :-

   i) N= 100   (A) =75   (B)=60   (AB) =40

   ii) N= 100   (A) =60   (B)=50   (AB) =20

3. Define   variable and attribute

4. From the following ultimate class frequencies, compute remaining

   Frequencies   (AB)=13 , (Aβ) = 20 , (αB)= 15 , (αβ) =9

5. If N=200, (A) =150, (B) =100, (AB)=80 find   (A) &(B).

6. Define positive and negative attribute.

7. Find the total no of class frequencies in case of two attribute.

8. Define the order of class.

9. If N=100 , (A) =550 , (B) =700 , &(AB) =300 find (αB) , (αβ).

10. 99% of the people who drink beer die before reaching age 80 years

    Therefore "drinking bear is bad for the longevity".

11. True or false

    i)   association is relation between two variable.

    ii)   attribute can we measure on the particular scale.

    iii)   The total no. of   ultimate classes in case of n attribute.

    iv)   The relation between height and weight is positive association.

12. Define attribute and give illustration.

13. Distinguish between a variable and attribute.

B) Questions for 4 marks

1. Distinguish between attribute and variable with example.

2. Show that co-efficient of association lies between -1 to 1

3. Define method of operation N. with example

4. Distinguish between association and correlation

5. Explain   i) positive and negative attribute

        ii) Ultimate class frequency

6. Define completely associated and dissociated of two attribute A& B

7. Define i) fundamental set of class frequency

        ii) Ultimate class frequency

8. Define attribute and give illustration

9. Define association and independence of two attribute

10. If $(A) = (B) = 4$, $N = 8$ obtain the co-efficient of association

   for each of the following cases

     i) $(AB) = 0$ ii) $(AB) = 1$

11. If the attribute A&B are independent then show that

    i)A and β ii) α and β are independent.

12. If Q is the Yule's coefficient between two attribute interpret the values

   $Q=0$, $Q=1$, $Q= -1$

13. In a certain interview there were 126 candidates of which 70 were

   boys, 36 candidates were successful among them 20 were boys obtained

   the co-efficient of association between and attribute boy.

14. Find the co-efficient of association between education standard and

   Employment from data and comment on result.

	Employed	Unemployed
Graduate	572	96
Non-Graduate	900	432

15. Out of 600 persons in a locality 150 attacked by cholera. In all

were inoculated against cholera of whom only 14 were attacked.

Comment on the effectiveness of inoculation in preventing cholera

As indicated by co-efficient of association

C). *Questions for 6 marks*

1. Explain the following
   a) Positive and negative attribute
   b) Order of the class
   c) Ultimate class frequencies
   d) independence of two attribute
   e) Positive and negative association
   f) Fundamental set of class frequency
2. Define the coefficient of association Q and interpret the case Q = 0,
   Q = 1
3. Show that coefficient of association Q lies between -1 and 1
4. If attribute A & B are independence then show that
   a) A and β b) α and B
5. Define method of operator N with example
6. Distinguish between association and correlation
7. Define completely associated and dissociated.
8. From the following ultimate class frequencies compute remaining
   frequency (AB) =13,(A β)=10 ,(αB)=9, (α β)=7.
9. The following are data on literacy and criminal tendency in a town
   Total population (in thousands)=224
   Total literates (in thousands) =40
   Literate criminals (in hundreds) =03
   Illiterate criminals (in hundreds)=43

   Find the association between literacy and criminal tendency and comment.

****