

Total No. of Printed Pages—12

1 SEM TDC STS M 1 (N/O)

2 0 2 1

(March)

STATISTICS

(Major)

Course : 101

(Descriptive Statistics)

*The figures in the margin indicate full marks
for the questions*

(New Course)

Full Marks : 80

Pass Marks : 24

Time : 3 hours

1. Choose the correct answer : 1×8=8

(a) Which one of the following is not quantitative data?

- (i) Weight of persons
- (ii) Sex of newborn babies
- (iii) Number of persons per family
- (iv) Percentage of marks in English

(b) The scale which does not possess any one of the three attributes such as magnitude, equal intervals and absolute zero point is a/an

- (i) ordinal scale
- (ii) nominal scale
- (iii) interval scale
- (iv) ratio scale

(c) The point of intersection of the 'less than' and the 'greater than' ogive corresponds to

- (i) mean
- (ii) median
- (iii) geometric mean
- (iv) None of the above

(d) The simple arithmetic mean of the first n natural numbers is

(i) $\frac{n(n+1)}{2}$

(ii) $\frac{n+1}{2}$

(iii) $\frac{2n+1}{3}$

(iv) None of the above

- (e) For any discrete distribution, standard deviation is
- (i) not less than mean deviation from mean
 - (ii) less than mean deviation from mean
 - (iii) equal to mean deviation from mean
 - (iv) None of the above
- (f) For a platykurtic curve
- (i) $\beta_2 > 3$
 - (ii) $\beta_2 < 3$
 - (iii) $\beta_2 = \infty$
 - (iv) $\beta_2 = 3$
- (g) When the correlation coefficient $r = \pm 1$, then the two regression lines
- (i) are perpendicular to each other
 - (ii) coincide
 - (iii) are parallel to each other
 - (iv) do not exist

(h) For two attributes A and B , which one of the following is not the condition for consistency of data?

(i) $(AB) \leq 0$

(ii) $(AB) \geq (A) + (B) - N$

(iii) $(AB) \leq (A)$

(iv) $(AB) \leq (B)$

2. (a) Distinguish between the following :

4

(i) Discrete data and Continuous data

(ii) Nominal scale and Interval scale of measurement

(b) What is statistical data? What are the points to be considered in classification of statistical data?

4

3. (a) Explain with examples, the various measurement scales used in statistics.

6

Or

(b) Name four different graphs used for presenting statistical data and discuss one of them in detail.

2+4=6

4. Prove that for any two positive quantities

$$AM \geq GM \geq HM$$

3

5. Answer any *three* of the following : $8 \times 3 = 24$

(a) When would you consider the median to be a better representative of the data than the arithmetic mean? Explain. Prove that for a frequency distribution, $\sum f_i (x_i - A)^2$ is minimum, when $A = \bar{x}$.

4+4=8

(b) What is 'standard deviation'? Explain its superiority over other measures of dispersion. Show that for any distribution, the standard deviation is not less than the mean deviation from the mean.

2+2+4=8

(c) Define moments. What is its use? Express first four central moments in terms of moments about the origin.

2+2+4=8

(d) What do you understand by skewness? Explain the methods of measuring skewness. Distinguish clearly, by giving figures, between positive and negative skewness.

2+4+2=8

6. (a) How can you use scatter diagram to obtain an idea of the extent and nature of correlation coefficient?

3

(b) Show that the lines of regression Y on X and X on Y pass through the point (\bar{X}, \bar{Y}) .

3

7. Answer any *two* of the following : 8×2=16

- (a) Prove that Spearman's rank correlation coefficient is given by

$$1 - \frac{6 \sum d_i^2}{n^3 - n}$$

where d_i denotes the difference between the ranks of the i th individuals. Also mention three advantages of Spearman's rank correlation over Karl Pearson's correlation coefficient. 5+3=8

- (b) Define (i) line of regression and (ii) regression coefficient. Find the equations to the lines of regression and show that the coefficient of correlation is the geometric mean of coefficients of regression. 2+3+3=8

- (c) Distinguish between multiple and partial correlations. Show that for a trivariate distribution, the multiple correlation can be expressed as

$$1 - R_{1(23)}^2 = (1 - r_{12}^2)(1 - r_{13.2}^2) \quad 3+5=8$$

8. (a) Explain the meaning of independence of two attributes A and B and give a criterion for their independence. If $\delta = (AB) - (AB)_0$, then with usual notations prove that

$$[(A) - (\alpha)][(B) - (\beta)] + 2N\delta = (AB)^2 + (\alpha\beta)^2 - (A\beta)^2 - (B\alpha)^2$$

1+3+5=9

Or

- (b) (i) Define coefficient of colligation and
(ii) Yule's coefficient of association. If $N = 1250$, $(A) = 780$, $(B) = 600$, $(AB) = 50$, then find the other frequencies. Is the data set consistent?
- 2+2+5=9

(Old Course)

Full Marks : 80

Pass Marks : 32

Time : 3 hours

1. State True or False :

1×8=8

- (a) In drawing histograms, the class intervals should be continuous.
- (b) A questionnaire is filled up by the respondent.
- (c) Cumulative frequency is not decreasing.
- (d) Mean lies between median and mode.
- (e) Mean deviation is least when taken from median.
- (f) The coefficient of skewness is zero when the distribution is symmetrical.
- (g) In case of perfect correlation, both the lines of regression are parallel to each other.
- (h) Square of Yule's coefficient of association cannot exceed 1.

2. Answer the following in brief : $4 \times 4 = 16$

(a) What is the difference between classification and tabulation? Explain. 4

(b) What do you mean by median of a distribution? How do you determine it graphically? $1 + 3 = 4$

(c) What do you understand by skewness? How is it measured? Write down the Bowley's formula for measuring skewness. $1 + 1 + 2 = 4$

(d) Show that correlation coefficient is independent of change of origin and scale. 4

3. Answer any *two* of the following : $7 \times 2 = 14$

(a) What is statistical data? Differentiate between primary data and secondary data. Define population and sample with reference to statistics. Give examples. $1 + 3 + 3 = 7$

(b) Describe various measurement scales used in statistics. 7

(c) How do you define chart, diagram and graph? What are the advantages of charts, diagrams and graphs? $3 + 4 = 7$

4. Answer any *three* of the following : $5 \times 3 = 15$

(a) What are the prerequisites for an ideal measure of central tendency? Prove that $AM \geq GM \geq HM$. $2+3=5$

(b) Explain the main difference between mean deviation and standard deviation. Prove that standard deviation is independent of change of origin and scale. $2+3=5$

(c) What are raw moments and central moments? Establish the relationship between central moments in terms of raw moments. $2+3=5$

(d) What is kurtosis? The standard deviation of a symmetrical distribution is 5. What must be the value of the fourth moments about mean in order that the distribution be (i) leptokurtic, (ii) mesokurtic and (iii) platykurtic? $2+3=5$

5. Answer any *three* of the following : $6 \times 3 = 18$

(a) Define Karl Pearson's coefficient of correlation. Show that $-1 \leq r \leq 1$, where r is the correlation coefficient. $2+4=6$

- (b) Write down the formula for Spearman's rank correlation coefficient $\rho(x, y)$. The rankings of ten students in two subjects A and B are as follows :

A	3	5	8	4	7	10	2	1	6	9
B	6	4	9	8	1	2	3	10	5	7

What is the coefficient of rank correlation? 2+4=6

- (c) Explain why there are two lines of regression in case of two variables. Derive the equation of regression line of Y on X. 2+4=6

- (d) Define partial correlation coefficient for three variables case. Show that

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}} \quad 2+4=6$$

6. (a) What are the conditions for consistency for two attributes A and B? Name different measures of association and discuss any one of them. What is the odd ratio? Explain with an example.

2+4+2+1=9

Or

(b) Define Yule's coefficient of association and the coefficient of colligation. Establish that

(i) $-1 \leq Q \leq 1$;

(ii) $Q = \frac{2Y}{1+Y^2}$;

where Y is the coefficient of colligation.
