3 (Sem–1/CBCS) STA HC 1

2022

STATISTICS

(Honours)

Paper : STA–HC–1016

(Descriptive Statistics)

Full Marks : 60

Time : Three hours

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

1. Answer the following questions as directed :
   (any seven) 1×7=7

   (a) The point of intersection of the 'less than type' and 'more than type' ogive corresponds to _____.

   (Fill in the blank)

   (b) What do you mean by price relative ?

(c) If $x_i/f_i (i=1,2,...,n)$ is a frequency

distribution and $u_i = \dfrac{x_i - a}{h}$, then which

one of the following is true

(i) $\bar{x} = \bar{u}$

(ii) $\bar{u} = h\bar{x}$

(iii) $\bar{x} - a = h\bar{u}$

(iv) $\bar{u} = a\bar{x} + h$

Symbols have their usual meaning.

(d) Define ordinal data.

(e) The value of mean deviation is minimum when deviations are taken w.r.t. _____. (Fill in the blank)

(f) The signs of the two regression coefficients are different.

(State true **or** false)

(g) For an asymmetrical distribution mean = 5, median = 4. Find the value of mode.

(h) Write Sheppard's correction for μ_4.

(i) Give the definition of variance for a frequency distribution $x_i/f_i (i=1,2,...,n)$.

(j) State the advantage of coefficient of variation over standard deviation.

(k) Write *two* demerits of geometric mean.

(l) If one of the regression coefficient is 1, the other must be

(i) greater than 1

(ii) lie between −1 and zero

(iii) less than 1

(iv) lie between −1 to +1

(Choose the correct option)

2. Answer *any four* of the following questions :

2×4=8

(a) Mention *two* limitations of statistics.

(b) Distinguish between frequency and non-frequency data.

(c) Define multiple correlation and partial correlation for a distribution involving the variables X_1, X_2 and X_3.

(d) What do you mean by dichotomous and manifold classification of attributes ?

(e) Prove that Fisher's index number satisfies factor reversal test.

(f) For what value of A the quantity

$$\sum_{i=1}^{n} f_i (x_i - A)^2$$

would be minimum ? Prove that.

(g) Define absolute moments and factorial moments.

(h) Which is the best measure of dispersion and why ?

3. Answer **any three** questions :   5×3=15

(a) Define raw moments and central moments. Derive the relationship between nth central moment and raw moments about the origin.   2+3=5

(b) Explain a histogram. How would you draw a histogram when the width of all classes are not equal ? State how a histogram is different from a bar diagram.   2+2+1=5

(c) Give *two* values x_1 and x_2, prove that

$$AM \geq GM \geq HM$$

Also show that

$$HM = (GM)^2 / AM   3+2=5$$

(d) (i) Prove that A.M. of the two regression coefficients is greater than the correlation coefficient.

(ii) Examine the consistency of the following data :

$$N = 1000, (A) = 600, (B) = 500,$$
$$(AB) = 50$$

(iii) When two attributes are said to be positively associated ?   2+2+1=5

(e) (i) Define CLIN. Interpret the result CLIN = 130.50

(ii) Mention *two* sources of secondary data.

(iii) What is a box plot ?   2+1+2=5

(f) How would you determine median graphically by using

(i) single ogive

(ii) both the ogives ?   2+3=5

(g) Write a note on skewness and kurtosis including different measures for them and relevant diagrams.

(h) Find the mean deviation from the mean and standard deviation of A.P. $a, a+d, a+2d,..., a+2nd$ and verify that the latter is greater than the former.   4+1=5

4. Answer *any three* questions :   10×3=30

(a) (i) Prove that $-1 \leq r_{XY} \leq +1$ 3

(ii) Are two uncorrelated variables essentially independent. If not, prove it with the help of an example.   2

(iii) Discuss the steps involved in the construction of wholesale price index numbers.   5

(b) (i) If for a random variable X the absolute moment of order k exists for ordinary $k = 1, 2, ..., n-1$, then the following quantities :

$$\beta_k^{2k} \leq \beta_{k-1}^k \cdot \beta_{k+1}^k \quad \text{and}$$

$$\beta_k^{1/k} \leq \beta_{k+1}^{1/k+1}$$

hold for $k = 1, 2, ..., n-1$, where β_k is the kth absolute moment about the origin.   4+1=5

(ii) Show that in a discrete series its deviations are small compared with mean μ so that $(x/M)^3$ and higher power of (x/M) are neglected, we have

$$G = M\left(1 - \frac{1}{2}\frac{\sigma^2}{M}\right)$$

where M is the arithmetic mean and G is the geometric mean. 5

(c) (i) Define the measures of association Q and Y and show that

$$Q = \frac{2Y}{1+Y^2} \qquad\qquad 2+3=5$$

(ii) Write the properties of multiple correlation coefficient. What is the significance of partical correlation coefficient in regression analysis ?

3+2=5

(d) (i) Describe the term 'deflation' in index number. 2

(ii) Find the angle between two lines of regression and interprete the result for $r=0$ and $r=\pm1$. 4+1=5

(iii) The regression equation of x on y is

$$3y - 5x + 180 = 0$$

Given that $\bar{y} = 4$, $\sigma_x^2 = \frac{9}{16}$ and $n = 4$. Find r and $\bar{x}$. 3

(e) (i) Show that Laspeyre's and Paasche's index numbers do not satisfy the time and factor reversal tests of consistency. 5

(ii) Interpret the meaning of the statement

$$byx = -0.53 \qquad\qquad 1$$

(iii) Distinguish between observational studies and controlled experiment with example.   4

(f) *(i)* Discuss the method of least squares for fitting a straight line $Y = a + bx$.   5

(ii) Find the regression line of Y on X. At which point this line intersects the regression line of X on Y.

4+1=5

(g) *(i)* Write a note on Sheppard's correction for moments. Define Pearson's β and γ coefficient.

3+2=5

(ii) Show that in usual notation

$$1 - R^2_{1.23} = \left(1 - r^2_{12}\right)\left(1 - r^2_{13.2}\right)$$ 3

(iii) Define partial correlation coefficient.

2

(h) *(i)* Write a note on scrutiny of data for internal consistency and detection of errors.   5

(ii) Give idea of cross-validation.   2

(iii) Write a note on chain index numbers.   3

———————